

BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 14

NUMBER 2

SPRING 1999

TABLE OF CONTENTS

SYMPOSIUM

THE LEGAL AND POLICY FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE: A PROGRESS REPORT

| | |
|--|-----|
| FOREWORD..... | 503 |
| By Kalama Lui-Kwan and Kurt Opsahl | |
| INTELLECTUAL PROPERTY AND THE DIGITAL ECONOMY: WHY THE ANTI- CIRCUMVENTION REGULATIONS NEED TO BE REVISED..... | 519 |
| By Pamela Samuelson | |
| COMMENTARY: BLACK HOLES OF INNOVATION IN THE SOFTWARE ARTS..... | 567 |
| By Mark A. Haynes | |
| AS MANY AS SIX IMPOSSIBLE PATENTS BEFORE BREAKFAST: PROPERTY RIGHTS FOR BUSINESS CONCEPTS AND PATENT SYSTEM REFORM..... | 577 |
| By Robert P. Merges | |
| OF GOVERNMENTS AND GOVERNANCE..... | 617 |
| By A. Michael Froomkin | |
| PROGRESSING TOWARDS A UNIFORM COMMERCIAL CODE FOR ELECTRONIC COMMERCE OR RACING TOWARDS NONUNIFORMITY?..... | 635 |
| By Maureen A. O'Rourke | |
| THE NEW MONEY..... | 659 |
| By Kerry Lynn Macintosh | |
| CLASH OF THE TITANS: REGULATING THE COMPETITION BETWEEN ESTABLISHED AND EMERGING ELECTRONIC PAYMENT SYSTEMS..... | 675 |
| By Jane Kaufman Winn | |

| | |
|--|-----|
| OLD AND NEW ISSUES IN THE TAXATION OF ELECTRONIC COMMERCE..... | 711 |
| By David L. Forst | |
| THE SPEED GAP: BROADBAND INFRASTRUCTURE AND ELECTRONIC COMMERCE..... | 721 |
| By Howard A. Shelanski | |
| STANDARDIZING GOVERNMENT STANDARD-SETTING POLICY FOR ELECTRONIC COMMERCE..... | 745 |
| By Mark A. Lemley | |
| THE LIMITS IN OPEN CODE: REGULATORY STANDARDS AND THE FUTURE OF THE NET .. | 759 |
| By Lawrence Lessig | |
| RESTORING AMERICANS' PRIVACY IN ELECTRONIC COMMERCE | 771 |
| By Joel R. Reidenberg | |

ARTICLE

| | |
|--|-----|
| DATABASE PROTECTION AT THE CROSSROADS: RECENT DEVELOPMENTS AND THEIR IMPACT ON SCIENCE AND TECHNOLOGY | 793 |
| By J. H. Reichman and Paul F. Uhlir | |

COMMENT

| | |
|---|-----|
| ELECTRONIC COMMERCE, HACKERS, AND THE SEARCH FOR LEGITIMACY: A REGULATORY PROPOSAL | 839 |
| By Michael Lee, Sean Pak, Tae Kim, David Lee, Aaron Schapiro, and Tamer Francis | |

FOREWORD: THE LEGAL AND POLICY FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE

By Kalama Lui-Kwan[†] and Kurt Opsahl[‡]

In July 1997, the Clinton Administration released its third major Internet initiative, a policy statement entitled *The Framework for Global Electronic Commerce*.¹ The Framework² outlines the Clinton Administration's

© 1999 Kalama Lui-Kwan and Kurt Opsahl.

[†] Symposium Editor, *Berkeley Technology Law Journal*; B.A., International Relations and China Studies, Boston University, 1995; J.D. candidate 2000, Boalt Hall School of Law, University of California, Berkeley; M.P.P. candidate 2000, John F. Kennedy School of Government, Harvard University.

[‡] Research Fellow; J.D. 1997, Boalt Hall School of Law, University of California, Berkeley.

The Conference was jointly organized by the Berkeley Center for Law & Technology and the *Berkeley Technology Law Journal*, and took place at the University of California at Berkeley on March 5-6, 1999. The authors wish to give special thanks to Professor Pamela Samuelson for her leading role in coordinating the event and for her support in developing this Symposium issue. The authors also wish to thank those who helped to organize the Conference, especially Larry Trask and John Sasson of the Berkeley Center for Law & Technology, Greg Papciak of the *Berkeley Technology Law Journal*, Pat Murphy of the Institute for Management, Innovation, and Organization, and Mark Lemley of the University of Texas at Austin. The authors are also grateful for the support of John Cioffi, for his excellent summary of the papers presented at the conference; Rachna Dhamija and David Chott for their work on a summary of the background issues; the Conference speakers, who contributed their insight and time to assure the high quality of the content of the Conference and this Symposium issue; the Conference volunteers, for helping to make the two-day event run without any logistical problems; and the editors of the *Berkeley Technology Law Journal* for the substantial support and time they have committed toward the success of this Symposium issue.

1. See WILLIAM J. CLINTON & ALBERT GORE, JR., A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE (1997), available at <<http://www.ecommerce.gov/framework.htm>> [hereinafter FRAMEWORK]. The first two initiatives were Ronald H. Brown, Secretary of Commerce, NATIONAL INFORMATION INFRASTRUCTURE: AN AGENDA FOR ACTION (Dec. 1993), available at <<http://metalab.unc.edu/nii/NII-Table-of-Contents.html>>, and Bruce Lehman, Patent and Trademark Office, REPORT OF WORKING GROUP ON INTELLECTUAL PROPERTY RIGHTS OF INFORMATION INFRASTRUCTURE TASK FORCE, INTELLECTUAL PROPERTY RIGHTS AND THE NATIONAL INFORMATION INFRASTRUCTURE (Sept. 1995), available at <<http://www.uspto.gov/web/offices/com/doc/ipnii/>> [hereinafter *White Paper*].

2. The Framework is often referred to as the Magaziner Report, after Ira Magaziner, former Senior Adviser to the President for Policy Development, who was princi-

strategy for facilitating the growth of electronic commerce and fostering business and consumer confidence in the use of electronic networks for commerce.³ The Framework presents five "principles" that are intended to "guide the development of the new digital economy."⁴ The principles are: (1) the private sector should lead, (2) governments should avoid undue restrictions on electronic commerce, (3) where government involvement is needed, its aim should be to support and enforce a predictable, minimalist, consistent, and simple legal environment for commerce; (4) governments should recognize the unique qualities of the Internet; and (5) electronic commerce over the Internet should be facilitated on a global basis.⁵

In addition, the Framework identifies nine areas where it makes recommendations to accomplish these principles: Tariffs and Taxation; Electronic Payment Systems; Uniform Commercial Code for Electronic Commerce; Intellectual Property Protection; Privacy; Security; Telecommunications Infrastructure and Information Technology; Content; and Technical Standards.⁶ Contemporaneously with the Framework, President Clinton issued thirteen directives to implement the report's recommendations.⁷ The directives are generally consistent with the philosophy that the private sector should take the lead role in developing self-regulation of electronic commerce markets.⁸

pally responsible for its creation. Magaziner has since resigned, and been replaced by Elliot Maxwell. See Jeri Clausing, *Magaziner, Head of U.S. Internet Policy, Plans to Resign*, N.Y. TIMES ON THE WEB (Nov. 6, 1998) <<http://www.nytimes.com/library/tech/98/11/cyber/articles/07magaziner.html>>; Jeri Clausing, *Commerce Dept.'s New Point Man on the Net*, N.Y. TIMES ON THE WEB (Dec. 11, 1998) <<http://www.nytimes.com/library/tech/98/12/cyber/articles/11maxwell.html>>.

3. The Framework was recently updated with the U.S. GOVERNMENT WORKING GROUP ON ELECTRONIC COMMERCE, FIRST ANNUAL REPORT (Nov. 1998), available at <<http://www.doc.gov/ecommerce/review.htm>> [hereinafter FIRST ANNUAL REPORT].

4. *Id.* at 10.

5. See FRAMEWORK, *supra* note 1, at 2-3.

6. See *id.* at 4-21.

7. See WILLIAM J. CLINTON, MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES, (July 1, 1997), available at <<http://www.ecommerce.gov/presiden.htm>> [hereinafter *Presidential Directive on Electronic Commerce*].

8. The U.S. Trade Representative is directed to ensure Internet commerce remains tariff-free, the Secretary of Commerce is directed to secure international intellectual property protections, and the Secretary of the Treasury is directed to discourage any new taxes from discriminating against Internet commerce. Other directives aim to protect privacy, ensure security, and promote electronic payment systems, including the creation of an online shopping system for the federal community. See *id.* See also Sandra Sobieraj,

While the Framework supports limited regulation in forms such as a uniform commercial code in cyberspace and the establishment of certain intellectual property norms, the main emphasis is to keep the government's role limited and enjoy the benefits of the Internet as "a global free-trade zone." It specifically implores the U.S. government and industry to work together to "adopt a non-regulatory, market-oriented approach to electronic commerce ... that facilitates the emergence of a transparent and predictable legal environment."⁹ One critique has been that the report expects the electronic commerce industry to take major initiatives, yet provides little incentive to undertake those initiatives.¹⁰ Consumer groups have also voiced skepticism about the private sector's ability to "self-regulate" collectively, especially on issues surrounding privacy and consumer protection.¹¹

The Berkeley Center for Law & Technology and the *Berkeley Technology Law Journal* convened a Conference two years after the Report was issued to assess the progress made on the Clinton Administration's proposals, and to examine the relationship between the explosive growth of electronic commerce and existing law and policy.¹² The Framework provided a focal point for the Conference because it is a reflection of the U.S. national policy on electronic commerce. The program included legal

Clinton Issues 'Hands Off' Policy on Internet Commerce, N.Y. TIMES ON THE WEB (July 2, 1997) <<http://www.nytimes.com/library/cyber/week/070297commerce.html>>.

9. FRAMEWORK, *supra* note 1, at 2.

10. For a survey of responses to the Framework, see Sobieraj, *supra* note 8, and Jeri Clausing, *Critics Question U.S. Policy on Electronic Commerce*, NY TIMES ON THE WEB (June 15, 1998) <<http://www.nytimes.com/library/tech/98/11/biztech/articles/30net.html>>. The primary incentive to self-regulate is the 'threat' of government regulation. See, e.g., FRAMEWORK, *supra* note 1, at 14 ("[I]f effective privacy protection cannot be provided [by self-regulation], we will reevaluate this policy.").

11. See Jeri Clausing, *Internet Commerce Study Stresses Self-Regulation*, N.Y. TIMES ON THE WEB (Nov. 30, 1998) <<http://www.nytimes.com/library/tech/98/11/biztech/articles/30net.html>>.

12. The full title of the Conference was *The Legal and Policy Framework for Global Electronic Commerce: A Progress Report*. For an overview of the Conference, see Berkeley Center for Law & Technology, *Electronic Commerce Conference* (visited Apr. 10, 1999) <<http://www.sims.berkeley.edu/bclt/ecom/>>. The authors would like to recognize the generous support of the following Conference sponsors: The School of Information Management and Systems, UC Berkeley; The Haas School of Business, UC Berkeley; BRIE (Berkeley Roundtable on the International Economy), UC Berkeley; Institute for Management, Innovation, and Organization, UC Berkeley; The IBM Institute for Advanced Commerce; The Fisher Center for Management & Information Technology; Cisco Systems, Inc.; The American Bar Association, Science & Technology Section.

scholars, technologists, government policy officials, and lawyers specializing in electronic commerce issues. To enable the insights from this Conference to be shared with a wider audience, the *Berkeley Technology Law Journal* agreed to publish this Symposium issue, which features selected papers that assess the past and future impact of existing laws, as well as papers that explore policy issues related to the Framework. This Symposium issue explores a broad range of topics discussed at the live Conference. Its design is to offer proposals for policy makers, concise discussions of certain sets of laws for industry officials, and forewarnings of policies and laws to come for legal practitioners.

I. GLOBAL STANDARDS FOR INTELLECTUAL PROPERTY RULES

In the digital economy, intellectual property rights will take a primary role, often forming the locus of value in a transaction. Understanding the growing importance of intellectual property rights, the Framework envisions a set of intellectual property rules that will enhance commerce in the global marketplace, yet remain “predictable, minimalist, consistent and simple.”¹³ Despite this laudable set of goals, and claims of substantial progress in the *First Annual Report*, it is not clear that the Clinton Administration has followed through with its own proposals.¹⁴ Three conference scholars addressed the issues surrounding intellectual property in the information economy, looking at copyright and patents in the new millennium.

Professor Pamela Samuelson, a Conference organizer and co-director of the Berkeley Center for Law & Technology, examines the recently enacted Digital Millennium Copyright Act (“DMCA”)¹⁵ and illustrates its inconsistency with the Framework’s principles.¹⁶ Samuelson’s article, *Intellectual Property and the Digital Economy: Why the Anti-Circumvention Regulations Need to Be Revised*, focuses on the anti-circumvention and anti-device provision of the DMCA. These provisions are designed to prevent engineers from developing or deploying technologies that circumvent technical protection mechanisms. The anti-circumvention provision pre-

13. FRAMEWORK, *supra* note 1, at 3.

14. See FIRST ANNUAL REPORT, *supra* note 3, at 10-13.

15. Pub. L. No. 105-304, 112 Stat. 2860 (1998).

16. See Pamela Samuelson, *Intellectual Property and the Digital Economy: Why the Anti-Circumvention Regulations Need to Be Revised*, 14 BERKELEY TECH. L.J. 519 (1999).

vents the act of circumvention while the anti-device provision outlaws the circumvention technologies themselves. Samuelson makes three main points regarding the provisions: there are legitimate reasons to circumvent technical protection mechanisms, the anti-device provisions are highly ambiguous and over-broad, and periodic reviews of the DMCA are needed to ensure that the Act's potential for mischief is not realized.

The DMCA's ban on circumvention lacks a general purpose "or other legitimate reasons" exception, instead choosing to codify seven carefully delineated exceptions, each responsive to a particular criticism. Samuelson explains how this inconsistency with the Framework's endorsement of simple and minimalist regulations is derived from both the poor judgment of the Clinton Administration and the extensive lobbying efforts of major copyright industries. The legislation became more complicated as the drafters included exceptions instead of reassessing the focus. Samuelson marshals a host of examples of legitimate circumvention, and suggests that the courts will narrow the reach of the DMCA's provisions if the Congress does not.

Building on the exceptions to the anti-circumvention provisions, Samuelson notes that the anti-device provision provides a broad ban without clarifying the legality of developing technologies that enable these exceptions. For example, under the Act, it is unclear if one is authorized to make devices that protect personal privacy, despite the DMCA's personal privacy exception. Indeed, Samuelson predicts that the breadth of the ban will engender a flood of litigation, as people question whether a particular technology fits in the provision's wide embrace. Until the courts or the Congress clarify the anti-device provisions, technologists may be deterred from developing legitimate technologies, thus slowing the pace of innovation. Finally, Samuelson finds the DMCA's call for a Library of Congress study too narrow. Rather, she argues for "a broader study to be undertaken of the impact of these regulations with an eye to recommending changes to remedy unintended harmful consequences they may be having."¹⁷

Silicon Valley patent attorney Mark Haynes criticizes the Framework's recommendation that steps be taken to ensure strong intellectual property protection.¹⁸ In *Black Holes of Innovation in the Software Arts*, Haynes argues that innovation cannot escape the gravity imposed by copyright law in certain areas. According to Haynes, copyright law creates

17. *Id.* at 564.

18. See Mark A. Haynes, *Commentary: Black Holes of Innovation in the Software Arts*, 14 BERKELEY TECH. L.J. 567 (1999).

isolated pockets of innovation by preventing anyone but the copyright holder from innovating. Without access to the copyrighted works, "Software engineers are constantly reinventing the wheel."¹⁹ This wastes valuable resources, and slows the pace of innovation, because the inventor must continue to go over old ground, rather than innovating.

Haynes looks at the Windows operating system, where Microsoft's copyright has helped to block the development of competing operating systems. Without competition, Haynes notes, the pace of OS innovation has been slower than that of hardware innovation.²⁰ Since 1995, Windows has undergone one major revision, integrating Microsoft Internet Explorer into the system for Windows 98, while the x86 microprocessors have undergone numerous substantial revisions.²¹ The difference lies in the system protection: Windows is protected by copyright, while Intel's x86 processors are protected by patents. Haynes calls for more explicit protection for reverse engineering of copyrighted software, like the approach taken in the Semiconductor Chip Protection Act of 1984,²² so the copyright system can follow the path of the patent system. This, Haynes opines, will allow the pace of software innovation to keep up with hardware innovation.²³

Professor Robert Merges' article, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts*²⁴ expands the examination of the patent system. Focusing on business method patents—which protect pure concepts, rather than technology²⁵—Merges argues that we should pay attention to the process by which patents are granted because the negative net effects are a potentially significant drag on the economy. Even though he says that it is practically impossible to determine the economic effect of business concept patents,²⁶ there are a number of policies that could be implemented to deal with the problem of issuing bad business concept patents. Merges, recognizing that drastic change is

19. *Id.* at 569.

20. *See id.* at 503.

21. *See id.*

22. 17 U.S.C. §§ 901-914 (1984).

23. *See* Haynes, *supra* note 18, at 575.

24. Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts*, 14 BERKELEY TECH. L.J. 577 (1999).

25. Until the recent *State Street* decision, patents on business concepts and other abstractions were simply not permitted. Merges explores the importance of business concept patents and, more specifically, whether they contribute any value in excess of their cost to society. *See, e.g.*, *Gottschalk v. Benson*, 409 U.S. 63 (1972); *State St. Bank & Trust Co., Inc. v. Signature Fin. Group, Inc.*, 149 F.3d 1368 (Fed. Cir. 1998).

26. *See* Merges, *supra* note 24, at 581.

unrealistic, outlines a number of modest proposals for reforming the patent office. Under one proposal, the patent office would be allowed to subcontract patent search and examination procedures to outside firms.²⁷ Merges also proposes that the patent office reduce the turnover rate of its examiners by increasing the quality of its examinations. Specifically, he suggests that senior examiners should be paid more, and that the patent office should increase expenditures for training their most junior people.²⁸ A third set of proposals involves reforming the patent examiner bonus system, which is "believed to skew incentives in favor of granting patents."²⁹ Merges offers two possible ways of reforming the bonus structure, each centering on creating disincentives to issue patents that are later determined to be invalid in court or reexaminations.³⁰

Merges concludes that trying to create a system within which examiners would never issue an invalid patent is unrealistic. The more reasonable goal of determining an acceptable "error rate," he argues, may be attainable by implementing these modest proposals.³¹ These proposals, Merges suggests, focus on the relationship between the patent office and the private sector in such a way as to allocate efficiently costs to determine patent validity.³²

II. GOVERNMENT REGULATION OF ELECTRONIC COMMERCE

Professor Michael Froomkin examines the issues surrounding Internet domain names, focusing on the creation and harmonization of governmental regulations of the domain name process. While lauding the initial declaration of principles in the Framework, Froomkin points out that the proposed action program "reveals a different view."³³ In *Of Governments and Governance*, Froomkin accuses the Clinton Administration of being "consumed by short-term policies and fail[ing] to grasp the consequences of the means proposed to achieve its short-term ends for long-term global governance."³⁴ Throughout the Framework, one can find calls for interna-

27. *See id.* at 604.

28. *See id.* at 608.

29. *Id.* at 609.

30. *See id.* at 609.

31. *See id.* at 615.

32. *See id.*

33. A. Michael Froomkin, *Of Governments and Governance*, 14 BERKELEY TECH. L.J. 617, at 620 (1999).

34. *Id.* at 620.

tional legal harmonization on electronic commerce issues. Froomkin looks to these edicts, and finds inconsistencies in the roles to be played by governments, industry, and international bodies in harmonizing the divergent rules. Harmonization is a challenging process, and subject to capture and other sub-optimal results. After a short discussion of the problems with the Article 2B process, Froomkin illustrates his point through an example with which he is deeply familiar: the World Intellectual Property Organization's domain name/trademark process.³⁵ WIPO is currently devising a process to address the issues created by the use of Internet domain names as marketing tools, as part of an international harmonization of domain name standards.³⁶ Noting the "dearth of consumer representatives, public interest groups, and citizen groups participating in the WIPO process," Froomkin argues that an elected government is a far better forum for providing notice to the public of the rules we expect them to obey.³⁷

Professor Maureen O'Rourke also sees challenges in implementing consistent global rules based on the Framework's principles. She looks at the conflict between the Framework's call for decentralized, market-led regulation of commercial transactions and the need for certainty in electronic commerce.³⁸ While she finds the high level objectives of the Framework sensible, the "devil, as always, has been in the details."³⁹ Even if there is international agreement on a broad framework, O'Rourke foresees the possibility of discord if the details are implemented at the local level. O'Rourke illustrates her point with the example of proposed Article 2B of the Uniform Commercial Code. Article 2B proposes to enunciate a model law of computer information transactions, and thereby validate the concept of shrinkwrap software licenses.⁴⁰ Despite some agreement that a

35. Froomkin sits on a Panel of Experts charged with assisting WIPO on its contribution to the Internet Domain Name Process, and has written extensively on the proposal.

36. For more information on the WIPO Internet Domain Name Process, see World Intellectual Property Organization, *Internet Domain Name Process* (visited Apr. 7, 1999) <<http://wipo2.wipo.int/>>.

37. See Froomkin, *supra* note 33, at 628.

38. See Maureen A. O'Rourke, *Progressing Towards a Uniform Commercial Code for Electronic Commerce or Racing Towards Nonuniformity?*, 14 BERKELEY TECH. L.J. 635 (1999).

39. *Id.* at 637.

40. See U.C.C. Article 2B (Feb. 1, 1999 Draft). Since that draft, the National Conference of Commissioners on Uniform State Laws ("NCCUSL") and the American Law Institute ("ALI") have decided to table U.C.C. Article 2B, and promulgate the legal rules for computer information transactions for adoption by states as the Uniform Computer Information Transactions Act. See ALI-NCCUSL joint press release, *NCCUSL to Promulgate Freestanding Uniform Computer Information Transactions Act—ALI and*

model law could be useful, Article 2B has been mired in controversy, and lacks a broad consensus on its controversial issues.⁴¹ Article 2B presents the essence of the conflict: a choice between uniform commercial standards that do not reflect commercial norms and the possibility of diverging legal standards that might slow the development of electronic commerce. O'Rourke calls for domestic and global conversations to resolve the "conflicting goals of flexibility and uniformity in the context of an overarching desire to encourage global electronic commerce."⁴²

Another issue critical to the growth of electronic commerce involves the development of international norms for electronic payment systems. This is the focus of Professor Kerry Lynn Macintosh's article *The New Money*.⁴³ While consumers today seem to prefer using credit cards over other electronic payment systems, Macintosh believes that it is "much too early to conclude Internet commerce can—or should—rely primarily on credit cards."⁴⁴ Instead, the Internet needs "global electronic currencies" that are privately issued, managed, and denominated.⁴⁵ The global marketplace, she argues, should not depend on the inflationary monetary policies and politics of sovereign nations.⁴⁶ Instead, the private sector should, as the Framework urges, take the lead by developing stable electronic currencies. To facilitate the growth of these global electronic currencies, Macintosh suggests that policy makers heed the call of the Framework to refrain from adopting "inflexible and highly prescriptive regulations and rules."⁴⁷ Generally, Macintosh believes that the proposals regarding electronic payment systems offer good policies, but are valuable only if the government takes steps that are consistent with the recommendations.

Professor Jane Winn continues the discussion of electronic payment systems in her article, *Clash of the Titans: Regulating the Competition*

NCCUSL Announce that Legal Rules for Computer Information Will Not Be Part of UCC (Apr. 7, 1999) <<http://www.2BGuide.com/docs/040799pr.html>>.

41. See, e.g., articles published in a dual symposium of the *California Law Review* and the *Berkeley Technology Law Journal* in January 1999 and December 1998 respectively.

42. O'Rourke, *supra* note 38, at 657.

43. Kerry Lynn Macintosh, *The New Money*, 14 BERKELEY TECH. L.J. 659 (1999).

44. *Id.* at 664.

45. *See id.*

46. *See id.* at 665.

47. *Id.* at 660 (citing FRAMEWORK, *supra* note 1, § 2).

*Between Established and Emerging Electronic Payment Systems.*⁴⁸ The new communications and information technologies, she says, have created conflicts over market dominance between new and established players.⁴⁹ While the new, electronic-based "Olympians" have had few successes,⁵⁰ the established, paper-based "Titans" still dominate the industry.⁵¹ Winn envisions a future in which the two sides collaborate to create a world in which new technology is used to access the existing infrastructure.⁵² The ultimate winner of the battle, if any, may be determined by whether the new entrants in the market become subject to the same regulations as the rest of the industry.⁵³

David Forst, a Silicon Valley tax attorney, outlines the difficulty of developing a tax regime for the digital environment in his article, *Old and New Issues in the Taxation of Electronic Commerce.*⁵⁴ The primary problem, he says, is that the Internet is still "a new medium whose full ramifications are not close to being understood."⁵⁵ The global nature of the Internet poses significant challenges to traditional notions of, among many other issues, jurisdiction, dispute resolution, and sovereignty. One question Forst examines within the context of electronic commerce is how the international community will fairly allocate online tax revenue as electronic commerce grows. The two options from which the world's nations have to choose are source-based taxation and residence-based taxation.⁵⁶ The trend, it seems, is toward a source-based form of taxation, which grants the country in which the enterprise earns the income the right to tax the profits of the enterprise.⁵⁷ The primary question now is how the international community will preserve the fiscal sovereignty of nations while equitably sharing the tax base from electronic commerce transactions.⁵⁸ This is particularly important, Forst notes, because, if a consistent interna-

48. Jane Kaufman Winn, *Clash of the Titans: Regulating the Competition Between Established and Emerging Electronic Payment Systems*, 14 BERKELEY TECH. L.J. 675 (1999).

49. *See id.* at 677.

50. *See id.* at 691-695.

51. *See id.* at 682.

52. *See id.* at 700.

53. *See id.* at 706.

54. David Forst, *Old and New Issues in the Taxation of Electronic Commerce*, 14 BERKELEY TECH. L.J. 711 (1999).

55. *Id.* at 711.

56. *See id.* at 712.

57. *See id.*

58. *See id.* at 716.

tional tax regime is to govern electronic commerce, it must be designed in such a way as to ensure that countries will not "believe that they are being denied their fair share of tax revenue."⁵⁹ Forst concludes that, while change in online tax regimes is inevitable, it will and should be incremental. In the meantime, policy makers should consider ways of relaxing barriers to countries taxing at the source by either expanding the use of consumption taxes or "liberaliz[ing] the permanent establishment principle in bilateral income tax treaties."⁶⁰

III. DEVELOPING ELECTRONIC COMMERCE INFRASTRUCTURE AND STANDARDS

In his article, *The Speed Gap: Broadband Infrastructure and Electronic Commerce*, Professor Howard Shelanski identifies a number of constraints underlying the Federal Communication Commission's advanced-services proceedings.⁶¹ He also suggests that policy makers should be wary of allowing these constraints to harm consumers or slow the development of affordable broadband services. Shelanski begins with an overview of the progress made in upgrading the telecommunications infrastructure to accommodate the demands of the national information economy. While fiber-optic cables have been deployed along freeways and main roads, they still have yet to reach individual consumers along neighborhood streets.⁶²

Most of the alternatives for extending broadband capacity to consumers are expensive and represent only one of many factors in determining consumer demand for online transactions. However, as costs fall and connection speeds increase, consumers will eventually be able to access more interactive forms of information, and will likely "engage not just in more transactions, but in more *kinds* of transactions as well."⁶³ For businesses, cheaper Internet access will make it easier to enter and survive in the electronic marketplace by expanding the market for online businesses, increasing the number of transactions made electronically, and expanding the types of transactions consumers can make online. For government, this means that FCC regulators will have to decide fairly soon between, on the

59. *Id.* at 715.

60. *Id.* at 716.

61. See Howard Shelanski, *The Speed Gap: Broadband Infrastructure and Electronic Commerce*, 14 BERKELEY TECH. L.J. 721 (1999).

62. See *id.* at 723.

63. *Id.* at 732.

one hand, the competitive benefits of opening the market to new entrants, and, on the other hand, the economies of scale and scope offered by incumbents that want to provide advanced services at prices lower than their competitors.⁶⁴ In other words, a choice will have to be made between the long-term benefits of competition and the short-term benefits of quality control and economies of scale.

Professor Mark Lemley finds that the Framework's proposals for technical standards have a lot of potential, and would be helpful if they were consistent with the government's actual approach. In his article, *Standardizing Government Standard-Setting Policy for Electronic Commerce*, Lemley explores the contradictions and gaps between the Framework and the standard-setting policies offered by the federal government.⁶⁵ Lemley argues first that electronic commerce definitely requires uniform standards, whether they are set by government mandate, industry organizations, or simply by a market "tip" in favor of one particular product.⁶⁶ The government's participation in setting these standards, however, runs against the policy document's strong position against government intervention. For example, while the Framework calls for the "development of a voluntary, market-driven key management infrastructure"⁶⁷ for encryption, the government has refused to allow companies to export or buy anything that fails to "meet its idea of a proper standard."⁶⁸

Lemley then explores the tension between the Framework's proposals and the government's policies. That is, while the Framework supports the strengthening of intellectual property rights in the digital environment, such rights are by definition "inimical to open standard setting."⁶⁹ Even standard setting organizations such as the Internet Engineering Task Force ("IETF"),⁷⁰ Lemley points out, have difficulty maintaining open standards in the presence of strong intellectual property rights in the standards themselves.⁷¹ Lemley argues that the way to resolve many of these tensions should start with government making a genuine effort to "get out of the way of private standard setting organizations that promote open stan-

64. See *id.* at 737.

65. See Mark Lemley, *Standardizing Government Standard-Setting Policy for Electronic Commerce*, 14 BERKELEY TECH. L.J. 745 (1999).

66. See *id.* at 747.

67. *Id.* at 748 (quoting FRAMEWORK, *supra* note 1, § 6).

68. *Id.* at 749.

69. *Id.* at 752.

70. For more information on the IETF, see Internet Engineering Task Force, *IETF Overview* (visited Apr. 8, 1999) <<http://www.ietf.org/overview.html>>.

71. See Lemley, *supra* note 65, at 750-51.

dards.”⁷² If the government chose to adopt a more progressive role, it could endorse interoperability and private sector initiatives by refusing to use or buy products that rely on closed proprietary standards.⁷³ The government, he concludes, has become involved when it has a stake in the outcome, even if the Framework says that the private sector should lead. Lemley argues that more than mere rhetoric is needed to deal with the questions posed by open and closed standards.

If the government continues to regulate cyberspace standards, despite the Framework’s rhetoric, how far can it go before infringing on liberties, and how should the private sector respond? In *The Limits in Open Code: Regulatory Standards and the Future of the Net*, Professor Lawrence Lessig argues for open code within applications as a method of preserving liberties in cyberspace.⁷⁴ Lessig notes that “code” is a kind of cyberspace law, and argues that we should examine the freedoms and constraints built into the code, as we would real space law.⁷⁵ This examination reveals that controlling the code behind Internet standards allows the government to regulate by technical means what might be difficult, if not impossible, to regulate through traditional legal rules. By stepping back a level, and regulating the code that regulates behavior, governments may be able to engender regulatory regimes that the cyber-libertarians thought impossible just a few years ago.⁷⁶

In order to preserve liberties in the face of possible regulation, Lessig looks to the open source movement. Closed code, where the source code is hidden, is subject to regulation to the extent that governments can control the originator. In contrast, open code, where the source code is made public, is far more challenging to regulate. Even if the law can mandate certain regimes to be reflected in the application, this does not insure that these regimes will be adopted. Lessig illustrates his point with the example of Netscape’s open-source browser, controlled by an organization known as Mozilla.⁷⁷ While Mozilla might be required to insert weak encryption into the public source code, this does not mean that an end user would not replace the weak encryption with strong encryption before compiling the application. This, Lessig asserts, shows that “the regulability of the appli-

72. *Id.* at 756.

73. *See id.* at 756.

74. *See* Lawrence Lessig, *The Limits in Open Code: Regulatory Standards and the Future of the Net*, 14 BERKELEY TECH. L.J. 759 (1999).

75. *See id.* at 763.

76. *See id.* at 761.

77. *See id.* at 766.

cation space turns in part on whether the application space is open.”⁷⁸ Lessig concludes with a note about the open source code movement. For many, Lessig notes, the virtue of making the source code public is in the efficiency and power that it enables. However, as he demonstrates in his article, there are greater issues at stake. Open code is not only valuable for utilitarian viability, but for the broader social values it supports and maintains.

Joel Reidenberg explains in his article, *Restoring Americans' Privacy in Electronic Commerce*, how a combination of technology and law can safeguard the privacy of electronic commerce consumers.⁷⁹ He argues that policymakers have adopted the theory that industry will eventually create adequate privacy protections for consumers on their own.⁸⁰ However, Reidenberg finds that companies have profited from the secret accumulation and sale of consumer information.⁸¹ Acxiom Corporation, for example, “sells information such as ethnic and religious affiliations, the type of car a person drives, and whether a person buys specialty clothing like particular types of underwear.”⁸² Even the industry’s self-regulation initiatives are less than reliable.

For its part, the Framework “ignores [the] incongruity” between American privacy policy and the global trend toward the establishment of comprehensive legal rights.⁸³ The enactment of comprehensive statutes, as in the case of Europe, has been slow. The European Directive took five years to assemble, and implementation within each nation was scheduled to take another three years.⁸⁴ Nevertheless, the Framework could have offered a few proposals to provide a new approach for privacy protection in the new environment of electronic commerce. Reidenberg offers three examples of what the Framework could have done. First, the United States should adopt the principles of the Organization for Economic Cooperation and Development, which offers a set of standards already recognized by American companies, and enact them into law.⁸⁵ Second, the government should act in a fashion that encourages technological devel-

78. See Lessig, *supra* note 74, at 767.

79. See Joel R. Reidenberg, *Restoring Americans' Privacy in Electronic Commerce*, 14 BERKELEY TECH. L.J. 771 (1999).

80. See *id.* at 774-775.

81. See *id.* at 776.

82. *Id.* at 4; see also Acxiom Direct Media, *Mailing Lists and More!* (visited Apr. 8, 1999) <<http://www.directmedia.com/>>.

83. *Id.* at 780.

84. See *id.* at 783.

85. See *id.* at 788.

opment, but only in such a way as to encourage “privacy protections rather than privacy intrusions.”⁸⁶ Third, the government should establish a U.S. Information Privacy Commission to serve “the tripartite role of consensus builder, privacy arbitrator and international advocate.”⁸⁷ Existing agencies and departments—such as the Department of Commerce, the State Department, and the White House Office of Management and Budget—are limited in the scope of their powers, and would not have as much authority to serve the three functions.⁸⁸

IV. CONCLUSION

One theme that emerged from the live Conference and echoes throughout this Symposium issue is the need for consistency and vision in federal laws governing electronic commerce. That is, policy makers should draft legislation that is both consistent with federally-established principles and, more importantly, visionary enough to balance the interests of consumers and the private sector over the long term. To its credit, the Framework recognizes the unique nature of the Internet, and focuses on the development of information as a commodity. Nevertheless, the policy document lacks the political strength necessary to align rhetoric with actual policies. While the Framework claims that the private sector should lead,⁸⁹ federal initiatives in encryption, for example, seem to suggest that the private sector should lead only when the government wants it to lead. The Framework also lacks the insight necessary to craft a legal and policy infrastructure that enhances both the long-term profitability of the private sector and provides adequate incentives and protections for innovators and consumers. The Digital Millennium Copyright Act, for example, may have allayed the concerns of copyright industry companies, which have argued that “fair use should not be an acceptable reason to ‘break’ a technical protection system used by copyright owners to protect their works.”⁹⁰ However, the legislation failed to alleviate the fears of Silicon Valley firms,⁹¹ librarians, and nonprofit groups⁹² that the anti-circumvention provisions would severely limit lawful use of and access to information. The actual policy also fails to conform to the Framework’s principle of sup-

86. *Id.* at 789.

87. *Id.* at 790.

88. *See id.*

89. *See, e.g.*, FRAMEWORK, *supra* note 1, at 2.

90. Samuelson, *supra* note 16, at 539.

91. *See id.* at 522.

92. *See id.* at 540.

porting and enforcing a “minimalist, consistent, and simple legal environment.”⁹³ Instead, it incorporates a short-sighted vision that maximizes near-term profits for a small set of intellectual property holders at the expense of long term economic growth. The future of the information economy requires the right balance of intellectual property rights with innovation, competition, and the free flow of information.⁹⁴

Neither this Symposium issue nor the actual Conference purport to reconcile all of the inconsistencies, or offer all of the necessary solutions, but together they form a two-pronged attempt to reveal legal and policy issues that require careful and balanced consideration. To be sure, the criticism offered at this Conference does in fact come with constructive and realistic suggestions for change. Still, the next step depends heavily on the integrity and vision of policy makers as they construct the rules that will guide the development of electronic commerce as we collectively lay the foundation for our global information society.

93. FRAMEWORK, *supra* note 1, at 3.

94. See William Landes & Richard Posner, *An Economic Analysis of Copyright Law*, 28 J. OF LEGAL STUD. 325, 326 (1989); Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 997-98 & n.32 (1997).

INTELLECTUAL PROPERTY AND THE DIGITAL ECONOMY: WHY THE ANTI-CIRCUMVENTION REGULATIONS NEED TO BE REVISED

By Pamela Samuelson[†]

ABSTRACT

The Digital Millennium Copyright Act of 1998 ("DMCA") prohibits the circumvention of technological protection measures used by copyright owners to control access to their works. It also bans devices whose primary purpose is to enable circumvention of technical protection systems. The Clinton administration proposed these anti-circumvention rules as implementations of U.S. obligations under the World Intellectual Property Organization Copyright Treaty. However, the DMCA's provisions are significantly broader than the treaty required. They violate the Administration's stated goal of only imposing "predictable, minimalist, consistent, and simple" regulations on the budding digital economy.

Although Congress heeded some concerns of digital economy firms by crafting certain exceptions to authorize legitimate circumvention, those exceptions are overly narrow and shortsighted. They should be supplemented by a more general "other legitimate purposes" exception. The DMCA's anti-device provisions are, moreover, overbroad and unclear, especially on the question whether it is legal to develop a technology necessary to engage in a privileged act of circumvention (e.g., a fair use). Either Congress or the courts will be forced to constrain the reach of the anti-device rules so as not to undermine Congressional intent to preserve fair uses and so as not to harm competition and innovation in the information technology sector. Finally, though the DMCA provides

© 1999 Pamela Samuelson.

[†] Professor of Information Management and of Law, University of California at Berkeley; Co-Director of the Berkeley Center for Law and Technology. This paper is an outgrowth of work initially done for an Emory Law School conference on the law of cyberspace held in February 1996. The draft article produced for that conference entitled *Technical Protection for Copyrighted Works* discussed a 1995 legislative proposal for regulating the circumvention of technical protection systems. I am deeply indebted to Benjamin Black who was my research assistant during preparation of this draft. He subsequently collaborated with me on a derivative work of that paper. Although that project was never completed, this article builds on the base of that collaboration. I am also grateful for comments on this draft from Hal Abelson, Jonathan Band, Yochai Benkler, Julie Cohen, Gideon Frieder, Joan Feigenbaum, Bob Glushko, Peter Huang, Laurel Jamtgaard, and Kurt Opsahl.

for a study of one class of potentially harmful impacts of the anti-circumvention rules, this study needs to be broadened to consider the full impact of this unprecedented legislation.

TABLE OF CONTENTS

| | |
|--|-----|
| I. INTRODUCTION | 520 |
| II. THE DIGITAL ECONOMY IS A HIGH GROWTH, HIGH POTENTIAL SECTOR WHOSE NEEDS DESERVE CAREFUL CONSIDERATION | 525 |
| III. THE WIPO COPYRIGHT TREATY IS GOOD FOR THE NEW ECONOMY | 528 |
| IV. DMCA'S OVERBROAD ANTI-CIRCUMVENTION PROVISIONS ARE NEITHER CONSISTENT WITH FRAMEWORK PRINCIPLES NOR GOOD FOR THE NEW ECONOMY | 534 |
| V. THE ENUMERATED EXCEPTIONS IN THE ACT-OF-CIRCUMVENTION BAN ARE UNDULY NARROW AND INCONSISTENT WITH FRAMEWORK PRINCIPLES | 537 |
| A. The Statutory Exceptions to the Circumvention Ban..... | 537 |
| B. Circumvention for Other Legitimate Reasons Should Be Privileged | 543 |
| VI. THE ANTI-DEVICE PROVISIONS SHOULD BE NARROWED BY LEGISLATIVE AMENDMENT OR JUDICIAL INTERPRETATION | 546 |
| VII. POLICYMAKERS SHOULD PERIODICALLY REVIEW BOTH THE ACT AND DEVICE PROVISIONS..... | 557 |
| VIII. CONCLUSION..... | 562 |

I. INTRODUCTION

The Clinton Administration's *Framework For Global Electronic Commerce* aims to promote the development of a vast global market in which electronic contracts will be made for delivery of electronic information products and services via digital networks which will be paid for with electronic currencies.¹ The Framework simultaneously encourages private investment and entrepreneurship, urges governments at all levels to act with restraint in considering regulations of the emerging digital economy, and argues for international cooperation in adopting consistent policies that will promote this commerce.² The Commerce Department's *First Annual Report* on the Framework initiative indicates that this initiative has

1. See WILLIAM J. CLINTON & ALBERT GORE, JR., A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE (1997), available at <<http://www.iitf.nist.gov/eleccomm/ecommm.htm>> [hereinafter FRAMEWORK].

2. See *id.* at 2-4.

met with some success.³ Passage of the Digital Millennium Copyright Act (“DMCA”)⁴ is among the successes claimed in this report.⁵

The Commerce Department may be correct in thinking that the interests of the digital economy will be furthered by widespread acceptance of the World Intellectual Property Organization (“WIPO”) Copyright Treaty⁶ in the international community.⁷ This treaty establishes several important international norms for applying copyright law in the digital environment.⁸ International consensus on these norms should aid the growth of the global digital economy.⁹ However, the DMCA was largely unnecessary to implement the WIPO Copyright Treaty because U.S. law already complied with all but one minor provision of that treaty.¹⁰

Although the WIPO Copyright Treaty requires countries to provide “adequate protection” against the circumvention of technical measures used by copyright owners to protect their works from infringement, the DMCA went far beyond treaty requirements in broadly outlawing acts of circumvention of access controls and technologies that have circumvention-enabling uses.¹¹

3. See U.S. GOV'T WORKING GROUP ON ELEC. COMMERCE, FIRST ANNUAL REPORT (1998), available at <<http://www.doc.gov/e-commerce/E-comm.pdf>> [hereinafter FIRST ANNUAL REPORT].

4. Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2360 (1998).

5. See FIRST ANNUAL REPORT, *supra* note 3, at 2.

6. See WIPO Copyright Treaty, adopted by the Diplomatic Conference on Dec. 20, 1996, WIPO Doc. CRNR/DC/94 (Dec. 23, 1996) [hereinafter WIPO Copyright Treaty]. There were actually two treaties concluded at this diplomatic conference. The other was the WIPO Performances and Phonograms Treaty, adopted by the Diplomatic Conference on Dec. 20, 1996, WIPO Doc. CRNR/DC/95 (Dec. 23, 1996). Because the U.S. protects the interests of producers and performers of phonograms largely through copyright law and because the phonograms treaty was not materially different in its requirements as regards issues covered in this article, the article will, for the sake of simplicity, focus on the WIPO Copyright Treaty provisions.

7. See generally Pamela Samuelson, *The U.S. Digital Agenda at WIPO*, 37 VA. J. INT'L L. 369 (1997) (discussing the negotiations leading to conclusion of the WIPO Copyright Treaty).

8. See *infra* notes 45-55 and accompanying text for a discussion of these norms.

9. See FIRST ANNUAL REPORT, *supra* note 3, at 10-11.

10. See, e.g., Pamela Samuelson, *Big Media Beaten Back*, WIRED, March 1997, at 64 (explaining that U.S. law was in compliance with almost all norms of the treaty). Only the treaty provision calling for protecting the integrity of rights management information needed legislative implementation in U.S. law. WIPO Copyright Treaty, *supra* note 7, art. 12; see also *infra* notes 56-64 and accompanying text.

11. WIPO Copyright Treaty, *supra* note 6, art. 11. The DMCA anti-circumvention provision can be found at 17 U.S.C.A. § 1201 (West Supp. 1999). See *infra* notes 66-70

The anti-circumvention rules in the DMCA do not match up well with the needs of the digital economy, or with the principles propounded in the Framework.¹² Although the *First Annual Report* praises the DMCA for the balance it embodies between copyright protection and access to information,¹³ this article will demonstrate that such balance as the DMCA contains is attributable to congressional foresight, not to the Clinton Administration.¹⁴ Indeed, for the past five years, the Administration has supported highly unbalanced digital copyright initiatives and has resisted most efforts to introduce more balance in these initiatives.¹⁵ With the enactment of the anti-circumvention provisions of the DMCA, the Administration may have had more success in achieving imbalance in digital copyright law than Congress may have realized.¹⁶

It would oversimplify the facts—although not by much—to say that the battle in Congress over the anti-circumvention provisions of the DMCA was a battle between Hollywood and Silicon Valley.¹⁷ Hollywood and its allies sought the strongest possible ban both on the act of circum-

and accompanying text for a discussion of why the treaty did not require the DMCA provisions.

12. See *infra* Part III for an articulation of these principles. See *infra* Parts V-VIII for an analysis of why these provisions may be harmful to digital economy interests.

13. See FIRST ANNUAL REPORT, *supra* note 3, at 14.

14. See *infra* Part V.

15. See U.S. DEP'T OF COMMERCE INFO. INFRASTRUCTURE TASK FORCE, INTELLECTUAL PROPERTY AND THE NATIONAL INFORMATION INFRASTRUCTURE: THE REPORT OF THE WORKING GROUP ON INTELLECTUAL PROPERTY RIGHTS (1995) [hereinafter *White Paper*]. Numerous articles have criticized this and an earlier draft report because of its imbalance heavily tilted in favor of publisher interests. See, e.g., Peter A. Jaszi, *Caught in the Net of Copyright*, 75 OR. L. REV. 299 (1996); Leslie Kurtz, *Copyright and the National Information Infrastructure*, 18 EUR. INTEL. PROP. REV. 120 (1996); Jessica Litman, *The Exclusive Right to Read*, 13 CARDOZO ARTS & ENT. L. 29 (1994); Charles R. McManis, *Taking TRIPS on the Information Superhighway: International Intellectual Property Protection and Emerging Computer Technology*, 41 VILL. L. REV. 207 (1996); Pamela Samuelson, *The Copyright Grab*, WIRED, Jan. 1996, at 134.

16. See *infra* Parts V-VII.

17. See, e.g., *WIPO Copyright Treaties Implementation Act; and Online Copyright Liability Limitation Act: Hearing on H.R. 2281 and H.R. 2280 Before the Subcomm. on Courts and Intellectual Property of the House Comm. on the Judiciary 105th Cong.* 78-82 (1997) [hereinafter *Judiciary Hearing*] (statement of Jack Valenti, President and CEO, Motion Picture Ass'n of America); *id.* at 256-65 (statement of Edward J. Black, President, Computer and Communications Industry Ass'n). It should be noted that the Business Software Alliance, whose principal member is Microsoft, supported Hollywood's preferred bill for reasons which may become apparent later in this article. See *infra* notes 180-186 and accompanying text. See also *Judiciary Hearing, supra*, at 68-77 (statement of Robert W. Holleyman II, President, Business Software Alliance).

venting a technical protection system used by copyright owners to protect their works and on technologies having circumvention-enabling uses.¹⁸ Silicon Valley firms and their allies opposed this broad legislation because of deleterious effects it would have on their ability to engage in lawful reverse engineering, computer security testing, and encryption research.¹⁹ They supported legislation to outlaw acts of circumvention engaged in for the purpose of infringing copyrights and would have supported narrowly drawn device legislation had the Congressional subcommittees principally responsible for formulating WIPO treaty implementation legislation been receptive to a narrower bill.²⁰ Silicon Valley and its allies warned of dire consequences if the overbroad anti-circumvention provisions Hollywood supported were adopted.²¹ Yet, by colorful use of high rhetoric and forceful lobbying, Hollywood and its allies were successful in persuading Congress to adopt the broad anti-circumvention legislation they favored, even if it is now subject to some specific exceptions that respond to some concerns raised by Silicon Valley firms and their allies in the legislative process.²²

Had the Administration sought to broker a fairer compromise between the interests of Hollywood and its allies and the interests of Silicon Valley and its allies, this process would almost certainly have produced better legislation than the anti-circumvention provisions of the DMCA. One would have thought, given the Framework's principles and the Administration's enthusiasm for the strong economic performance of the infor-

18. See, e.g., *Judiciary Hearing*, *supra* note 17, at 78-82 (statement of Jack Valenti); *id.* at 204-12 (statement of Allan R. Adler, Vice President for legal and governmental affairs, Ass'n of American Publishers).

19. See *infra* notes 87-94 and accompanying text. Other groups opposed to the broad anti-circumvention legislation of H.R. 2281 included librarians and educators. See *infra* notes 117-120 and accompanying text.

20. The Digital Future Coalition—whose members include the Computer & Communications Industry Association, among other high tech industry groups—endorsed H.R. 3048, 105th Cong. (1997), which proposed such a narrow circumvention provision. See *Introduction of the Digital Era Copyright Enhancement Act*, 55 BNA PAT., TRADEMARK & COPYRIGHT J. 68, 70-71 (1997) (describing the anti-circumvention provision of H.R. 3048). See also *Judiciary Hearing*, *supra* note 17, at 256-65 (statement of Edward J. Black) (critical of the Administration's anti-circumvention proposal); *id.* at 249-56 (statement of Chris Byrne, Director of Intellectual Property, Silicon Graphics, Inc., on behalf of the Info. Tech. Indus. Council) (critical of H.R. 2281).

21. See, e.g., *Judiciary Hearing*, *supra* note 17, at 260 (prepared statement of Edward J. Black); see also *id.* at 154-55 (prepared statement of Prof. Robert L. Oakley, Georgetown University Law Center).

22. See *infra* Part V.

mation technology sector, that the Administration would have taken a more balanced position on these issues.²³ One can call the DMCA's anti-circumvention provisions many things, but one cannot honestly speak of them as "predictable, minimalist, consistent, and simple" components of a legal environment for electronic commerce, as the Framework principles would suggest they should be.²⁴

This article will make three main points about the anti-circumvention rules in the DMCA. First, there are far more legitimate reasons to circumvent a technical protection system than the DMCA's act-of-circumvention provision expressly recognizes.²⁵ This provision should be amended to provide a general purpose "or other legitimate purposes" provision to avert judicial contortions in interpreting the statute. Second, the anti-device provisions of the DMCA are highly ambiguous and overbroad, raising questions about whether Congress understood the potential for these provisions to undermine circumvention privileges built into the act-of-circumvention prohibition.²⁶ The anti-device provisions of DMCA should be clarified and a more minimalist approach taken to the regulation of technologies with circumvention-enabling uses so that the ambiguity and overbreadth of the existing provisions will not cause harm to innovation and competition in the information technology sector. Third, periodic reviews of the impact of the anti-circumvention provisions of the DMCA as a whole should be undertaken.²⁷ Given how broad the anti-circumvention rules are, given their unprecedented character, and given the potential for harmful consequences from these rules, Congress should authorize a far broader study of the impact of these provisions than the DMCA presently contemplates. It should also heed proposals for change to the anti-circumvention provisions recommended in such studies.

23. See *infra* Part III.

24. See FRAMEWORK, *supra* note 1, at 3. For further criticism of the DMCA's anti-circumvention provisions on constitutional grounds, see Yochai Benkler, *Free As the Air To Common Use: First Amendment Constraints on the Enclosure of the Public Domain*, 74 N.Y.U. L. REV. 354 (1999).

25. See *infra* Part VI.

26. See *infra* Part VII.

27. See *infra* Part VIII.

II. THE DIGITAL ECONOMY IS A HIGH GROWTH, HIGH POTENTIAL SECTOR WHOSE NEEDS DESERVE CAREFUL CONSIDERATION

An April 1998 report, *The Emerging Digital Economy*, published by the U.S. Department of Commerce begins with the following observations:

During the past few years, the United States economy has performed beyond most expectations. A shrinking budget deficit, low interest rates, a stable macroeconomic environment, expanding international trade with fewer barriers, and effective private sector management are all credited with playing a role in this healthy economic performance.

Many observers believe advances in information technology ("IT"), driven by the growth of the Internet, have also contributed to creating this healthier-than-expected economy.

In recent testimony to Congress, Federal Reserve Board Chairman Alan Greenspan noted, "our nation has been experiencing a higher growth rate of productivity—output per hour—worked in recent years. The dramatic improvements in computing power and communication and information technology appear to have been a major force behind this beneficial trend."²⁸

This report indicates that the IT sector of the U.S. economy—which includes the computer hardware, software, networking and telecommunications industries—now constitutes an estimated 8.2 per cent of the gross domestic product, close to twice its share of GDP as compared with a decade or so before.²⁹ The IT sector, moreover, accounts for more than one-quarter of the real economic growth in the American economy.³⁰ Approximately 45 per cent of current expenditures on business equipment are investments in IT products and services.³¹ It is no wonder, then, that the collective capitalization of five major firms in this sector—Microsoft, Intel, Compaq, Dell, and Cisco Systems—has grown from \$12 billion in 1987 to \$588 billion in 1997, nearly a fifty-fold increase in only a dec-

28. U.S. DEP'T OF COMMERCE, SECRETARIAT ON ELEC. COMMERCE, *THE EMERGING DIGITAL ECONOMY 1* (1998) [hereinafter *EMERGING DIGITAL ECONOMY*].

29. *See id.* at 4.

30. *See id.* at 6.

31. *See id.*

ade.³² Perhaps somewhat more wondrous are the astonishing market capitalizations of relatively new Internet firms, such as Amazon.com, Yahoo!, and E*Trade. These valuations reflect the market's belief in the high growth potential of these players in the digital economy, even if their earnings so far might seem to belie this.³³ It is, of course, important to realize that the IT sector is not the only component of the digital economy.³⁴ It is, however, a significant part of that economy, and it is also the enabler of growth in other parts of the digital economy, as vendors of products and services of both tangible and intangible kinds make use of digital networks to offer their wares to a global market.³⁵ Especially as electronic commerce via the Internet and the World Wide Web expands, the IT sector is likely to experience further explosive growth.³⁶

The Emerging Digital Economy report continues along the path set by the Administration's early policy document, *The Framework for Global Electronic Commerce*, in seeking to foster the growth potential of the digital economy.³⁷ Both documents recognize that "[g]overnments can have a profound effect on the growth of commerce on the Internet. By their actions, they can facilitate electronic trade or inhibit it. Knowing when to act and—at least as important—when not to act, will be crucial to the development of electronic commerce."³⁸ One of the signal achievements of the Framework was the promulgation of five principles that were supposed to guide U.S. as well as other governmental action on policy initiatives on electronic commerce:

- 1) The private sector should lead.
- 2) Governments should avoid undue restrictions on electronic commerce.

32. *See id.* Of course, it is fair to observe that some of this growth has occurred by virtue of acquisitions of other substantial firms, such as Compaq's acquisition of Digital Equipment Corp.

33. *See, e.g.,* James J. Cramer, *TulipMania.com? Despite their soaring prices, the best Internet stocks are still bargains. Here's how to pick 'em*, TIME, Aug. 3, 1998, at 77; *see generally* Steve Mott, *Where Eagles Soar: Making Sense of Internet Valuations*, BUSINESS 2.0, Nov. 1998.

34. *See* EMERGING DIGITAL ECONOMY, *supra* note 28, chs. 4-5 (discussing digital economy sectors).

35. *See id.*

36. *See id.*

37. *See id.* at 50-51.

38. FRAMEWORK, *supra* note 1, at 2; EMERGING DIGITAL ECONOMY, *supra* note 28, at 50-51.

- 3) Where government involvement is needed, its aim should be to support and enforce a predictable, minimalist, consistent, and simple legal environment for commerce.
- 4) Governments should recognize the unique qualities of the Internet.
- 5) Electronic commerce over the Internet should be facilitated on a global basis.³⁹

The *First Annual Report* of the U.S. Working Group on Electronic Commerce offers evidence that the Framework's policy objectives are being achieved.⁴⁰

As laudable as the Framework's principles are, it should be said that the Clinton Administration has been somewhat erratic in following them. The Administration has a good record in promoting minimalist tax and customs policies.⁴¹ However, it has been widely criticized by the IT/digital economy sector for not following these principles in the security/encryption policy area and in the content policy area, owing to the Administration's support for the Clipper Chip and the Communications Decency Act.⁴² In the legislative struggle leading up to adoption of the DMCA, the Administration deviated from these principles once again in heeding the desires of established copyright industries to reconstruct the legal infrastructure of the digital environment so that it would accommodate their preferences. These industries insisted that this restructuring was necessary to protect them from the grave threat of piracy posed in the digital environment.⁴³ Many significant players in the existing digital economy counseled against this restructuring.⁴⁴ The Administration should, of course, have considered the interests and concerns of Hollywood and other copyright industry groups in its consideration of an appropriate digital copyright policy initiative. However, the Administration might have done more to consider the interests of those already partici-

39. FRAMEWORK, *supra* note 1, at 2-3.

40. *See id.* at iii-v.

41. *See id.* at iii, 7 (mentioning passage of the Internet Tax Freedom Act); *see also id.* at 12 (discussing foreign tax initiatives).

42. *See, e.g.*, ESTHER DYSON, RELEASE 2.0 (1997).

43. *See Judiciary Hearing, supra* note 17, at 79-80 (prepared statement of Jack Valenti).

44. *See id.* (testimony of Edward J. Black; testimony of Chris Byrne); *see also The WIPO Copyright Treaties Implementation Act: Hearing on H.R. 2281 Before the Subcomm. on Telecomm., Trade, & Consumer Protection of the House Comm. on Commerce, 105th Cong.* (1998) [hereinafter *Commerce Hearing*].

pating in the digital economy in its policy formation on these issues, particularly since its preferred policy so clearly violated the principles that the Administration had asserted it would follow.

III. THE WIPO COPYRIGHT TREATY IS GOOD FOR THE NEW ECONOMY

The WIPO Copyright Treaty established several norms about applying copyright law in the digital environment.⁴⁵ They include:

- 1) copyright owners should have an exclusive right to control the making of copies of their works in digital form,⁴⁶
- 2) copyright owners should have an exclusive right to control the communication of their works to the public,⁴⁷
- 3) countries can continue to apply existing exceptions and limitations, such as fair use, as appropriate in the digital environment, and can even create new exceptions and limitations appropriate to the digital environment,⁴⁸

45. See WIPO Copyright Treaty, *supra* note 7. See also Samuelson, *supra* note 7 (discussing the digital agenda WIPO treaty provisions).

46. There was an explicit provision on the reproduction right in the draft treaty initially considered at WIPO. See Basic Proposal For the Substantive Provisions of the Treaty On Certain Questions Concerning the Protection of Literary and Artistic Works To Be Considered at the Diplomatic Conference, WIPO Doc. CRNR/DC/4, art. 7(1) (Aug. 30, 1996). However, this provision did not attract consensus because of its inclusion of temporary reproductions, which was highly controversial. See Samuelson, *supra* note 7, at 382-90. Instead, the diplomatic conference agreed on certain statements of interpretation of the treaty which included a provision on the reproduction right. See Agreed Statements Concerning the WIPO Copyright Treaty, adopted by the Diplomatic Conference on Dec. 20, 1996, WIPO Doc. CRNR/DC/96 at 1 (Dec. 23, 1996) [hereinafter Agreed Statements]. For a discussion of the tortured history of the draft treaty provision, the Agreed Statements, and what they mean, see Samuelson, *supra* note 7, at 382-92.

47. See WIPO Copyright Treaty, *supra* note 6, art. 8. While the United States does not have an exclusive right of communication in its copyright law, see 17 U.S.C. § 106 (1994) (exclusive rights provisions), its public performance and distribution rights are substantively equivalent to this right. See *id.*; Samuelson, *supra* note 7, at 392-98 (discussing negotiations concerning digital communications).

48. See Agreed Statements, *supra* note 46, at 2. This agreed statement was in striking contrast to the proposed treaty language and proposed comments on exceptions and limitations to copyright in the draft treaty considered at the WIPO diplomatic conference. See Samuelson, *supra* note 7, at 398-409 (discussing the draft and final provisions on fair use and other exceptions). Although the White Paper had expressed doubts about the vi-

- 4) merely providing facilities for the communication of works should not be a basis for infringement liability,⁴⁹
- 5) it should be illegal to tamper with copyright management information insofar as this would facilitate or conceal infringement in the digital environment,⁵⁰ and
- 6) countries should have “adequate legal protection and effective legal remedies against the circumvention of effective technological measures” used by copyright owners to protect their works from infringing uses.⁵¹

To the extent that uncertainties about how copyright law should apply in the digital environment were impeding the growth of a global market in electronic intellectual property products,⁵² there was reason to be optimis-

ability of fair use in the digital environment, the Clinton Administration was ultimately persuaded that the WIPO Copyright Treaty should contain a more positive statement about fair use in the digital environment. See White Paper, *supra* note 15, at 82; Samuelson, *supra* note 7, at 406.

49. See Agreed Statements, *supra* note 46, at 2. This issue had been highly contentious, both in the U.S. and at the diplomatic conference, because the Clinton Administration supported holding online service providers strictly liable for infringing acts of their users. See White Paper, *supra* note 15, at 114-24; Samuelson, *supra* note 7, at 385-88 (discussing controversy at diplomatic conference). The DMCA included a provision substantially limiting on online service provider liability. See 17 U.S.C.A. § 512 (West Supp. 1999).

50. See WIPO Copyright Treaty, *supra* note 7, art. 12. For a discussion of the history and meaning of this provision, see Samuelson, *supra* note 7, at 415-18.

51. See WIPO Copyright Treaty, *supra* note 7, art. 11. The draft treaty considered at WIPO included a provision quite similar to the anti-circumvention provision endorsed by the Clinton Administration in the White Paper which sought to outlaw technologies, the primary purpose or effect of which was to circumvent technical protection measures. The draft treaty provision, like the White Paper's proposed anti-circumvention regulation, was highly controversial within the United States and even more so at the diplomatic conference. Many delegations expressed concern about the impact of such regulations on fair uses and public domain information. As a consequence, the final treaty included only a very general norm on anti-circumvention. See Samuelson, *supra* note 7, at 409-15.

52. Other factors besides uncertainties about the application of copyright law in the digital environment may be responsible for the slower-than-anticipated growth in the market for digital versions of copyrighted works. See, e.g., Pamela Samuelson, *Authors' Rights in Cyberspace: Are New International Rules Needed?*, FIRST MONDAY (Oct. 1996), available at <<http://www.firstmonday.dk/issues/issue4/samuelson/index.html>>. However, there is a better case for such uncertainties being an impediment on an international scale than in the United States. That U.S. copyright law protects authors against unauthorized digital reproductions of their works has been clear since 1979. See NATIONAL COMM'N ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS, FINAL

tic that conclusion of this treaty would remove these blockages and allow e-commerce to flourish.⁵³ These norms are as “predictable, minimalist, consistent, and simple” components of a legal environment for commerce as one could expect copyright professionals to devise.⁵⁴ Thus, the WIPO treaty itself established norms compatible with Framework principles and with the needs of the digital economy. That nearly one hundred sixty nations signed this treaty indicated a strong consensus that digital works should be given appropriate protection on an international scale.⁵⁵ This was very good news for U.S. digital economy industries.

The WIPO treaty digital copyright norms were, however, mostly old news for U.S. law.⁵⁶ Its cases had already recognized the rights of authors to control digital reproductions of their works,⁵⁷ as well as to control digital transmissions of their works to the public.⁵⁸ Courts had invoked fair use in a number of digital copyright cases,⁵⁹ and had refused to hold online service providers liable for infringing activities of users about which the providers had no knowledge.⁶⁰ Because of the substantial accord between the WIPO treaty norms and existing U.S. law, the Clinton Administration initially considered whether the WIPO Copyright Treaty might even be sent to the Senate for ratification “clean” of implementing legislation.⁶¹ This would have avoided the kind of protracted legislative battle that oc-

REPORT (1979). In some countries, however, this was not as clear. Insofar as the WIPO Copyright Treaty clarified this on an international basis, it did contribute to the legal infrastructure for global e-commerce. See Samuelson, *supra* note 7, at 382-85 (discussing lack of clarity about the reproduction right in the digital environment).

53. See, e.g., FIRST ANNUAL REPORT, *supra* note 3, at 13-14.

54. FRAMEWORK, *supra* note 1, at 3.

55. See List of Participants, WIPO Doc. No. CRNR/DC/INF.2 (Dec. 20, 1996).

56. The WIPO Copyright Treaty, as finally concluded, was actually far more consistent with U.S. copyright law than the draft treaty with which the negotiations had begun (and which was substantially based on proposals by U.S. officials). See Samuelson, *supra* note 7, at 434-37.

57. See, e.g., *Sega Enterprises, Ltd. v. MAPHIA*, 857 F. Supp. 679 (N.D. Cal. 1994).

58. See, e.g., *Playboy Enterprises, Inc. v. Frena*, 839 F. Supp. 1552 (M.D. Fla. 1993).

59. See, e.g., *Lewis Galoob Toys, Inc. v. Nintendo of America*, 964 F.2d 965 (9th Cir. 1992) (software enabling temporary changes in the play of Nintendo games held fair use).

60. See, e.g., *Religious Tech. Center v. Netcom Online Comm. Corp.*, 907 F. Supp. 1361 (N.D. Cal. 1995) (online service provider should not be held strictly liable for user infringement of which it had no knowledge).

61. See *Clinton Administration Is Undecided On Implementing Steps For WIPO Treaties*, 53 BNA PAT., TRADEMARK & COPYRIGHT J. 241 (1997).

curred when Congress considered the Administration's White Paper legislation in 1996.⁶² Eventually, the Administration decided that implementing legislation was necessary for the U.S. to comply with the WIPO treaty provision requiring protection for the integrity of copyright management information.⁶³ The DMCA implementation of this norm, which closely tracks the treaty language, was uncontroversial during the legislative process.⁶⁴

The U.S. could have asserted that its law already complied with the WIPO treaty's anti-circumvention norm.⁶⁵ This norm was, after all, very

62. See Samuelson, *supra* note 7, at 427-32 (arguing that U.S. efforts at WIPO conference were aimed at bypassing contention over domestic legislative proposals).

63. See WIPO Copyright Treaty, *supra* note 7, art. 12. Had this treaty defined the term "rights management information" ("RMI") only as "information which identifies the work, the author of the work, the owner of any right in the work," the U.S. could have relied on section 43(a) of the Lanham Act to assert that it was in compliance with the norms of this Article as well. See Julie E. Cohen, *Some Reflections on Copyright Management Systems and Laws Designed to Protect Them*, 12 BERKELEY TECH. L.J. 161, 169 n.31. However, the treaty defines RMI as including "information about the terms and conditions of use of the work, or any numbers or codes that represent such information...." WIPO Copyright Treaty, *supra* note 6, art. 12. Section 43(a) would not seem to cover misrepresentations of this sort. See 15 U.S.C. § 1125(a) (1994); see also Cohen, *supra*, at 169 n.31. In addition, it appears that some technical amendments to U.S. law were necessary to change the terminology about which foreign nationals could claim rights under U.S. law. See Section-by-Section Analysis of H.R. 2281 As Passed By the United States House of Representatives on August 4, 1998, 105th Cong., at 3-4 (1998) [hereinafter House Manager's Report].

64. See 17 U.S.C.A. § 1202 (West Supp. 1999). Concerns had earlier been expressed that copyright management systems might be intrusive on privacy interests of users. See, e.g., Julie E. Cohen, *The Right to Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace*, 28 CONN. L. REV. 981 (1996). In response to concerns of this sort, the legislative history of DMCA makes clear that copyright management information ("CMI") does not include digital information used to track or monitor usage of copyrighted works: "It would be inconsistent with the purpose and construction of this bill and contrary to the protection of privacy to include tracking and usage information within the definition of CMI." House Manager's Report, *supra* note 63, at 20.

65. It is far more plausible that the U.S. is in compliance with the WIPO treaty anti-circumvention norm than that it is in compliance with the moral rights provision of the Berne Convention, which is one of the minimum standard rules required of Berne Union members. See Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, art. 6bis (Paris Text, 1971, amended 1979), reprinted in 1 BASIC DOCUMENTS OF INT'L ECON. L. (CCH) 715 (1994). See also Jessica Litman, *The Tales That Article 2B Tells*, 13 BERKELEY TECH. L.J. 931, 932 (1998) (discussing the U.S. rationale for claiming to be in compliance with the Berne Convention's moral rights provision, and expressing skepticism about the adequacy of this rationale). See also Jonathan Band & Taro

general in character and provided treaty signatories with considerable latitude in implementation. Moreover, anti-circumvention legislation was new enough to many national intellectual property systems, and certainly to international law, to mean that there was no standard by which to judge how to instantiate the norm. The U.S. could have pointed to a number of statutes and judicial decisions that establish anti-circumvention norms.⁶⁶ With U.S. copyright industries thriving in the current legal environment, it would have been fair to conclude that copyright owners already were adequately protected by the law.⁶⁷ Even many of those who favor use of technical systems to protect digital copyrighted works have expressed skepticism about the need for or appropriateness of anti-circumvention regulations, at least at this stage.⁶⁸ Let content producers build their technical fences, advised one prominent information economist, but do not legislatively reinforce those fences until experience proves the existence of one or more abuses in need of a specific cure.⁶⁹ However, the political reality and legislative dynamics of the WIPO Copyright Treaty implementation process were such that some sort of anti-circumvention provision appeared to be a necessary part of the bill.

Even if a reasoned assessment of U.S. law might have led policymakers to conclude that some additional anti-circumvention legislation was necessary or desirable, one would have thought that the Administration would have supported a “predictable, minimalist, consistent, and simple”

Isshiki, *The New Anti-Circumvention Provision in the Copyright Act: A Flawed First Step*, 3 CYBERSPACE LAW. 2 (1999) (explaining that the DMCA's anti-circumvention regulations were not required for compliance with the WIPO Copyright Treaty).

66. See White Paper, *supra* note 15, at 232-34 (discussing statutes); *Sega Enterprises, Ltd. v. MAPHIA*, 857 F. Supp. 679 (N.D. Cal. 1994) (finding copyright liability for providing tools to enable game software to be removed from disks and posted on the Internet).

67. See, e.g., *Judiciary Hearing*, *supra* note 17, at 78 (statement of Jack Valenti) (citing \$60 billion in annual U.S. revenues from international sales of intellectual property and naming copyright industry as single greatest contributor to U.S. economy); Motion Picture Ass'n of America Research Dep't, *MPAA 1998 U.S. Economic Review* (visited Apr. 22, 1999) <<http://www.mpa.org/useconomicreview/1998/index.htm>> (demonstrating steadily increasing U.S. box office receipts between 1991 and 1998).

68. See, e.g., Tom W. Bell, *Fair Use vs. Fared Use: The Impact of Automated Rights Management on Copyright's Fair Use Doctrine*, 76 N.C. L. REV. 557, 561-62 (1998); David Friedman, *In Defense of Private Orderings*, 13 BERKELEY TECH. L.J. 1151, 1163-64 n.31 (1998).

69. See Ejan Mackaay, *The Economics of Emergent Property Rights on the Internet*, in *THE FUTURE OF COPYRIGHT IN A DIGITAL ENVIRONMENT* 13, 21 (P. Bernt Hugenholtz ed., 1996). “It is this restraint,” says MacKaay, “that guards us from sliding into rent-seeking.” *Id.* at 22.

legal rule, as its Framework principles call for. The Administration might have, for example, proposed to make it illegal to circumvent a technical protection system for purposes of engaging in or enabling copyright infringement. This, after all, was the danger that was said to give rise to the call for anti-circumvention regulations in the first place. Silicon Valley Representative Tom Campbell proposed such an approach in his alternative bill.⁷⁰ If this same assessment caused policymakers to decide there was also a need for some regulation of circumvention technologies to promote electronic commerce, then a “predictable, minimalist, consistent, and simple” legal rule would have been to outlaw making or distributing a technology intentionally designed or produced to enable copyright infringement.⁷¹ Many “digital economy” firms and organizations supported the first of these proposals,⁷² and they would likely have supported the second if it had ever had a chance of being taken seriously.

Clinton Administration officials, bowing to the wishes of Hollywood and its allies, opted instead to support an unpredictable, overbroad, and maximalist set of anti-circumvention regulations. During Congressional consideration of these provisions, these regulations became complex and inconsistent for reasons that will become evident in later sections of this article.⁷³ It was, in short, not the needs of the digital economy that drove adoption of the anti-circumvention provisions in the DMCA. Rather, what drove the debate was high rhetoric, exaggerated claims, and power politics from representatives of certain established but frightened copyright indus-

70. See H.R. 3048, 105th Cong. § 8 (1997). Northern Virginia Representative Rick Boucher (whose district includes America Online) cosponsored this bill.

71. This was how most previous regulations of circumvention technologies had been framed. See, e.g., Thomas C. Vinje, *A Brave New World of Technical Protection Systems*, 8 EUR. INTELL. PROP. REV. 431 (1996).

72. See *supra* note 20.

73. The anti-circumvention regulations are one of a number of amendments to the Copyright Act of 1976 that are contributing to its becoming increasingly unreadable. See, e.g., 17 U.S.C. § 104A (1994) (restoration of copyright in foreign works that had fallen into the public domain for lack of compliance with U.S. formality rules in effect until 1989). This is not to say that the 1976 Act was a model of comprehensibility in all respects. See, e.g., 17 U.S.C. §§ 111-112 (1994) (effective Jan. 1, 1978) (exceptions permitting passive retransmission of broadcast signals by cable systems and ephemeral recordings during broadcast transmission). However, these incomprehensible provisions had at least been negotiated by affected industry sectors who understood what the provisions meant, even if virtually no one else could comprehend them. In contrast, the restoration of foreign copyright and the new anti-circumvention regulations affect a broad range of industries. This makes the incomprehensibility of the provisions more troublesome.

tries. These groups seem to believe they are so important to America that they should be allowed to control every facet of what Americans do with digital information.⁷⁴ They also seem to think they are entitled to control the design and manufacture of all information technologies that can process digital information.⁷⁵ The DMCA caters to their interests far more than to the interests of the innovative information technology sector or of the public.

IV. DMCA'S OVERBROAD ANTI-CIRCUMVENTION PROVISIONS ARE NEITHER CONSISTENT WITH FRAMEWORK PRINCIPLES NOR GOOD FOR THE NEW ECONOMY

There are three principal rules in the final DMCA's anti-circumvention provision. The first focuses on the act of circumvention. Section 1201(a)(1)(A) generally outlaws the act of circumventing "a technological measure that effectively controls access to a work protected under this title."⁷⁶ This rule will, however, not take effect for two years from enactment, in part to allow time for a study to be conducted of the potential impact of this norm on noninfringing uses of copyrighted works.⁷⁷ When it does come into force, it will be subject to seven complex exceptions that will be discussed below in Part V.A.⁷⁸

Section 1201 also contains two "anti-device" provisions. Sections 1201(a)(2) and 1201(b)(1) both regulate technologies with circumvention-enabling capabilities. The former focuses on devices that circumvent "a technological measure that effectively controls access to a [copyrighted] work" (access controls).⁷⁹ The latter relates to devices that circumvent the "protection afforded by a technological measure that effectively protects a

74. See Samuelson, *supra* note 15 (discussing the copyright maximalist agenda the Clinton Administration has supported).

75. The potential for broad anti-circumvention regulations to give copyright owners power to control the design of consumer electronics products was recognized in Geneva. See John Browning, *Africa 1, Hollywood 0*, WIRED, March 1997, at 61, 186 ("Japan and other Asian nations were up in arms about proposals that would effectively have turned the consumer electronics industry into a branch of publishing."). Indeed, some unnoticed provisions of the DMCA will require the makers of consumer videotape recorders to build in anti-copying technology in subsequent models. See 17 U.S.C.A. § 1201(k).

76. 17 U.S.C.A. § 1201(a)(1)(A).

77. See *id.*; *infra* notes 208-210 and accompanying text.

78. See *id.* § 1201(d)-(j), discussed *infra* notes 98-135 and accompanying text.

79. *Id.* § 1201(a)(2); see also *id.* § 1201(a)(3) (defining the phrases "circumvent a technological measure" and "effectively controls access to a work").

right of a copyright owner ... in a work or a portion thereof" (e.g., copy controls).⁸⁰ In each case, section 1201 states that "[n]o person shall manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof"⁸¹ if it (1) "is primarily designed or produced for the purpose of circumventing,"⁸² (2) "has only limited commercially significant purpose or use other than to circumvent,"⁸³ or (3) "is marketed by that person or another acting in concert with that person with that person's knowledge for use in circumventing"⁸⁴ the technological measure or the protection it affords. The anti-device rules have a narrower range of exceptions than does the act-of-circumvention ban.⁸⁵

One would have to admit that the act-of-circumvention rule initially sought by the Administration was simpler, and at least in this respect, more consistent with the Framework's principles than the DMCA as enacted. The original proposal would have outlawed circumventions of technical protection systems except when done for legitimate law enforcement or intelligence purposes.⁸⁶ However, representatives of major information technology firms and organizations brought to Congress's attention that this norm would interfere with many legitimate activities.⁸⁷ It would, for example, have outlawed encryption research and computer security testing, even though these activities are critical to achieving many of the ob-

80. *Id.* § 1201 (b)(1); *see also id.* § 1201(b)(2) (defining the terms "circumvent protection afforded by a technological measure" and "effectively protects a right of a copyright owner under this title").

81. *Id.* § 1201(a)(2), (b)(1).

82. *Id.* § 1201(a)(2)(A), (b)(1)(A). There is no definition of "primarily designed or produced" in the statute; nor are any criteria for determining it provided in the statute.

83. *Id.* § 1201(a)(2)(B), (b)(1)(B). This subsection may be the broadest and most dangerous of the three conditions because it would seem to put at risk "freeware" or "shareware" programs that, by their very nature, have no commercial uses. MIT Professor Hal Abelson has informed me that he expressed his reservations about this subsection to Rep. Barney Frank who serves on the House Intellectual Property Subcommittee. Prof. Abelson said that this provision should outlaw technologies having "only limited legitimate uses." He reports that Rep. Frank agreed with this assessment. Yet the final provision retains the "limited commercial purposes" construction with which it began. Email correspondence with Hal Abelson (Feb. 28, 1999) (on file with author).

84. 17 U.S.C.A. § 1201(a)(2)(C), (b)(1)(C).

85. *See id.* § 1201(g)(4), (j)(4).

86. *See* H.R. 2281 § 1201, 105th Cong. (1997) (as introduced in the House of Representatives on July 29, 1997), *reprinted in* 54 BNA PAT., TRADEMARK & COPYRIGHT J. 270 (1997).

87. *See, e.g., Judiciary Hearing, supra* note 17, at 256-61 (statement of Edward J. Black).

jectives of the digital economy.⁸⁸ As Congress came to recognize that there were a number of legitimate reasons to circumvent technical protection systems, the bill slowly accreted exceptions that made the bill more complicated but less harmful to growth of the digital economy.⁸⁹

These same firms and organizations, in alliance with major consumer electronics firms, were also critical of the Administration's preferred anti-device provisions.⁹⁰ However, these digital economy groups exhausted their political capital on getting critical exceptions to the act-of-circumvention ban⁹¹ and on establishing that they had no affirmative duty to build their technologies to respond to technical protection systems, but only a duty to refrain from actively undermining them.⁹² They took some comfort in statements by Congressional supporters of a limited interpretation of the anti-device norms indicating that Congress meant for the anti-device provisions to apply to "black boxes" that are expressly intended to facilitate circumvention.⁹³ Still, the digital economy sector remains understandably concerned about the potential for overbroad application of

88. See Letter from Dr. Charles Brownstein, Chair of the Public Policy Committee of the U.S. Chapter of the Association for Computing Machinery, to Rep. Thomas J. Bliley, Chairman of the House Commerce Committee (Sept. 29, 1998) (on file with author) (expressing concern about impact of broad anti-circumvention regulations on computer security research). See also FRAMEWORK, *supra* note 1, § 6 (emphasizing the importance of computer security to the growth of global economic commerce).

89. See *infra* Part V.

90. See *Commerce Hearing*, *supra* note 44, at 32-33 (prepared statement of Chris Byrne, Director of Intellectual Property, Silicon Graphics, Inc., on behalf of Info. Tech. Indus. Council); *id.* at 28-30 (statement of Jonathan Callas, Chief Technology Officer, Network Assocs., Inc.); *id.* at 58-63 (statement of Seth Greenstein, Esq., on behalf of the Digital Media Ass'n); *id.* at 46-49 (statement of Walter H. Hinton, Vice President, Storage Tech. Corp., on behalf of the Computer and Communications Indus. Ass'n); *id.* at 18-27 (statement of Gary J. Shapiro, Chairman, Home Recording Rights Coalition, and President, Consumer Elecs. Mfrs. Ass'n).

91. See 17 U.S.C.A. § 1201(f), (g), and (j).

92. See *id.* § 1201(c)(3); 144 CONG. REC. H7093, H7095 (daily ed. Aug. 4, 1998) (statement of Rep. Bliley).

93. See *id.* at H7094-95 ("This provision is not aimed at products that are capable of commercially significant noninfringing uses...."). See also *id.* at H7097 ("[I]t is not enough for the primary effect of the device to be circumvention. It, therefore, excludes legitimate multipurpose devices...."); House Manager's Report, *supra* note 63, at 9 ("[Section 1201(a)(2)] is carefully drafted to target 'black boxes' and to ensure that legitimate multipurpose devices can continue to be made and sold."); *infra* note 192 and accompanying text.

the anti-circumvention and anti-device norms, and recent developments suggest that there is reason for this concern.⁹⁴

Although Administration officials admitted in Congressional testimony that its preferred legislation went beyond what the WIPO Copyright Treaty required, it argued for this broader rule in part to set a standard that would help the U.S. persuade other countries to pass similarly strong rules.⁹⁵ Proponents of the Administration's preferred anti-circumvention regulations scoffed at arguments made by an alliance of consumer electronics firms and by representatives of the computer and software industries about the harm that broad anti-circumvention regulations would do in this industry.⁹⁶ They also dismissed as specious arguments made by library and educational groups about threats to fair use and the public domain arising from broad anti-circumvention regulations.⁹⁷

V. THE ENUMERATED EXCEPTIONS IN THE ACT-OF-CIRCUMVENTION BAN ARE UNDULY NARROW AND INCONSISTENT WITH FRAMEWORK PRINCIPLES

A. The Statutory Exceptions to the Circumvention Ban

The DMCA ban on the act of circumventing technical protection systems is subject to seven very specific exceptions,⁹⁸ as well as being qualified by several other subsections.⁹⁹ In addition, it is subject to a two-year moratorium during which the Librarian of Congress is supposed to study the potential impact of the anti-circumvention ban on noninfringing uses of copyrighted works which may lead to further limitations on the act-of-circumvention rule.¹⁰⁰ While several of these exceptions and limitations respond to the gravest of concerns expressed by digital economy firms,¹⁰¹

94. See *infra* notes 193-195 and accompanying text.

95. See, e.g., *House Subcommittee Holds Hearings on WIPO Treaty Bills, OSP Liability*, 54 BNA PAT., TRADEMARK & COPYRIGHT J. 414 (1997).

96. See, e.g., *Judiciary Hearing*, *supra* note 17, at 204-12 (statement of Allan Adler).

97. See, e.g., *id.* at 229, 235-36 (testimony of Michael K. Kirk, executive director, American Intellectual Property Law Ass'n).

98. See 17 U.S.C.A. § 1201(d)-(i) (West Supp. 1999).

99. See *id.* § 1201(c)(1)-(4).

100. See *id.* § 1201(a)(1)(A)-(C).

101. See *id.* § 1201(f) (reverse engineering exception), 1201(g) (encryption research), and 1201(j) (computer security testing). See also *Judiciary Hearing*, *supra* note 17, at 260-61 (prepared statement of Edward J. Black) (expressing concern about reverse engi-

they are still too narrowly crafted, as examples given below will reveal.¹⁰² Congress should have adopted a provision enabling courts to exempt acts of circumvention engaged in for other legitimate purposes. Courts interpreting section 1201 may either be forced to find liability in some situations in which it would be inappropriate to impose it or to stretch existing limitations. Congress may eventually need to revise this provision to recognize a broader range of exceptions.

The structure of the final DMCA anti-circumvention provision and its complexity resulted from the maximalist position with which the Administration and its major copyright industry allies began the legislative struggle. Only when IT industry groups were able to identify particularized situations in which circumvention was appropriate was there any legislative "give" on the issue, and then only to the extent of that identified situation.¹⁰³ As noted above, Clinton Administration officials initially sought an almost unlimited ban of circumvention activities.¹⁰⁴ The only exception to the circumvention ban in the Administration's favored legislation was an authorization of circumvention of technical protection systems for legitimate law enforcement, intelligence, and other governmental purposes.¹⁰⁵ Without this exception, suspected Mafia bosses and terrorists, oddly enough, might have been able to challenge attempted law enforcement or intelligence agency decryptions of their records or communications under section 1201(a)(1).¹⁰⁶

The Administration's preferred bill also provided that nothing in section 1201 would "affect rights, remedies, limitations, or defenses to copy-

neering); *Commerce Hearing*, *supra* note 44, at 29-30 (prepared statement of Jonathan Callas) (expressing concern about encryption and security research).

102. *See infra* Part V.B.

103. *See supra* note 101.

104. *See* Band & Isshiki, *supra* note 65 (indicating that Patent and Trademark Office (PTO) officials had initially sought to outlaw circumvention of copy controls, as well as of access controls, and that lobbying by library and educational groups had persuaded Commerce Department officials to drop this provision of the PTO's preferred bill).

105. *See* H.R. 2281 § 1201(e), 105th Cong. (1997) (as introduced in the House of Representatives on July 29, 1997). The DMCA version of § 1201 has such a provision, although it has been expanded to enable government agencies to test the vulnerabilities of their computer systems or networks. *See* 17 U.S.C.A. § 1201(e) (West Supp. 1999).

106. Virtually all such records would likely embody a modicum of originality that would enable these actors to claim copyright protection in fixations of these records. If these persons used technical protection systems to prevent unauthorized access to these records, any act of the government to circumvent such systems would, strictly speaking, run afoul of § 1201(a)(1).

right infringement, including fair use, under this title.”¹⁰⁷ This seemed to recognize that circumventing a technical protection system for purposes of engaging in fair use or other noninfringing acts would be lawful, although it did not directly say so.¹⁰⁸ Some representatives of major copyright industries who testified at a Congressional hearing on this legislation expressed the view that fair use should not be an acceptable reason to “break” a technical protection system used by copyright owners to protect their works.¹⁰⁹ Allan Adler, testifying on behalf of the Association of American Publishers, for example, stated that “the fair use doctrine has never given anyone a right to break other laws for the stated purpose of exercising the fair use privilege. Fair use doesn’t allow you to break into a locked library in order to make ‘fair use’ copies of the books in it, or steal newspapers from a vending machine in order to copy articles and share them with a friend.”¹¹⁰ The “breaking and entering” metaphor for circumvention activities swayed some influential Congressmen in the debate over anti-circumvention regulations.¹¹¹

Courts should distinguish between circumvention aimed at getting unauthorized access to a work and circumvention aimed at making noninfringing uses of a lawfully obtained copy.¹¹² Section 1201(a)(1) is aimed at the former, not the latter. Fair use, for example, would provide a poor

107. H.R. 2281 § 1201(d) (as introduced in the House of Representatives on July 29, 1997). *See* 17 U.S.C.A. § 1201(c)(1).

108. An extremely narrow interpretation of the provision might suggest that fair use could be raised as a defense to an infringement claim based on activities engaged in after a circumvention had taken place (e.g., reproducing a portion of the work for fair use purposes), even if the act of circumvention itself would not be excused. *See Judiciary Hearing, supra* note 17, at 235-36 (testimony of Michael K. Kirk).

109. *See also* White Paper, *supra* note 15, at 231 (indicating that copyright owners have no obligation to make their works available in a form that will enable fair uses to be made of them).

110. *Judiciary Hearing, supra* note 17, at 208 (prepared statement of Allan Adler). This same speaker went on to say that “[t]he Declaration of Independence is in the public domain, but there is nothing wrong with the National Archives keeping it in a vault and punishing anyone who tries to break through security to get hold of that copy.” *Id.*

111. *See* House Manager’s Report, *supra* note 63, at 5 (characterizing circumvention to get unauthorized access as “the electronic equivalent to breaking into a locked room to obtain a copy of a book”). *But see, e.g.,* Friedman, *supra* note 68, at 1163 n.31 (arguing against the treatment of technologies capable of circumventing technical protection systems as “the digital equivalent of burglar’s tools”).

112. *See* Cohen, *supra* note 63, at 174-76 (discussing lawful circumvention); *see also* Julie E. Cohen, *Copyright and The Jurisprudence of Self-Help*, 13 BERKELEY TECH. L.J. 1089, 1142 n.200 (1998) (finding in copyright’s fair use doctrine an affirmative right to “hack” technical protection systems to make fair uses).

excuse for breaking into a computer system in order to get access to a work one wished to parody. However, if one had already lawfully acquired a copy of the work, and it was necessary to bypass a technical protection system to make fair use of that copy, this would appear to be lawful under section 1201(a)(1) and (c)(1).¹¹³ Take, for example, an act of circumvention performed by Geoffery Nunberg, a friend of mine who works for Xerox's Palo Alto Research Center. He was an expert witness in a lawsuit which successfully challenged the Washington Redskins' trademark on the ground that the word "redskins" is scandalous or disparaging.¹¹⁴ Nunberg decided it was necessary to take a clip from an old Western movie to demonstrate derogatory uses of the term in context. It was necessary for him to defeat a technical protection system adopted by the owner of the copyright in this movie in order to make the clip for this purpose. If section 1201(c)(1)'s preservation of fair use and other defenses to infringement are to be given their plain meaning, it would seem that this sort of circumvention should be permissible.¹¹⁵ Thus, if the clip from the movie qualifies as a fair use, the act of circumvention may be privileged under section 1201(c)(1).¹¹⁶

Although this section's apparent preservation of fair use was important, it did not satisfy library and nonprofit groups who expressed substantial concern about the impact that the anti-circumvention provisions would have on public access to information.¹¹⁷ The only additional concession that the House Subcommittee on Intellectual Property thought should be made to concerns expressed by these groups was to create a special "shopping privilege" for them. This exception, which was included in the final DMCA, enables nonprofit library and educational institutions to circum-

113. See 144 CONG. REC. H7097 (daily ed. Aug. 4, 1998) (letter from Rep. Howard Coble to Rep. Rick Boucher) (indicating an intent to distinguish between circumvention to get unauthorized access to a work and circumvention to make fair uses).

114. See *Harjo v. Pro-Football, Inc.*, 45 U.S.P.Q.2d (BNA) 1789 (1998); 15 U.S.C. § 1052(a) (1994) (excluding scandalous and disparaging matter from trademark protection); See also "*Redskins*" Mark is Cancelled as Disparaging to Native Americans, BNA PAT., TRADEMARK & COPYRIGHT LAW DAILY (Apr. 12, 1999).

115. See, e.g., 144 CONG. REC. H7093 (daily ed. Aug. 4, 1998) (statement of Rep. Bliley) (indicating that the Commerce Committee understood the legislation to enable consumers to "exercise their historical fair use rights"); see also *id.* at H7097 (letter from Rep. Coble to Rep. Boucher).

116. But see *infra* notes 157-162 and accompanying text for a discussion about whether this person's development of a technology enabling him to defeat the technical protection system would be similarly privileged.

117. See, e.g., *Commerce Hearing*, *supra* note 44, at 64-66 (statement of Prof. Robert L. Oakley).

vent technical protection systems to “make a good faith determination of whether to acquire a copy” of the work.¹¹⁸ Librarians and educators do not see much value in this provision because vendors of technically protected copyrighted works will generally have incentives to allow librarians and educators to have sufficient access to make acquisition decisions.¹¹⁹ Their broader concerns about the impact of anti-circumvention regulations on noninfringing uses fell on deaf ears in both the House and Senate Subcommittees on Intellectual Property.¹²⁰

Computer and software industry groups were initially unsuccessful in persuading Congress to create additional exceptions to the anti-circumvention rules and other changes to the anti-circumvention regulations to make them less harmful to legitimate activities in these industries.¹²¹ Not until the full Senate Judiciary Committee and the House Commerce Committee undertook their reviews of the legislation were concerns of these industry groups heeded. Out of the Senate Committee emerged three significant changes to the DMCA. The first was creation of a new exception to enable circumvention of technical protection systems for purposes of enabling a software developer to achieve interoperability among computer programs.¹²² The second was a provision clarifying that equipment manufacturers were under no obligation to specially design their products to respond to any particular technical measure used by those providing content for this equipment.¹²³ The third was a provision indicating that section 1201 was not intended to broaden contributory or vicarious copyright liability.¹²⁴

An interesting twist in the saga leading up to adoption of the DMCA was the House Commerce Committee’s decision to exercise jurisdiction

118. See 17 U.S.C.A. § 1201(d) (West Supp. 1999).

119. See *infra* notes 151-156 and accompanying text, concerning whether the shopping privilege could be undermined by the lack of available tools to enable this circumvention.

120. See, e.g., *Judiciary Hearing, supra* note 17, at 148-56 (statement of Robert L. Oakley); *id.* at 64-68 (statement of M.R.C. Greenwood, chancellor of the University of California, Santa Cruz) (expressing concerns about the impact of technical protection systems on noninfringing uses of protected works—concerns the “shopping privilege” does not address).

121. See, e.g., *id.* at 256-65 (statement of Edward J. Black) (expressing concern about the impact of the anti-circumvention provisions for achieving interoperability among computer programs).

122. See 17 U.S.C.A. § 1201(f) (West Supp. 1999).

123. See *id.* § 1201(c)(3).

124. See *id.* § 1201(c)(2).

over part of the digital copyright legislation.¹²⁵ Its review led to several other significant changes to the bill. Some of these responded to concerns expressed by digital economy firms; others responded to concerns expressed by library, educational, and other nonprofit groups.¹²⁶ The Commerce version of the bill added a new exception to enable encryption research and the development of encryption-research tools.¹²⁷ It also created two consumer-oriented exceptions, one to enable parents to circumvent access controls when necessary to protect their children from accessing harmful material on the Internet, and the other to enable circumvention to protect personal privacy.¹²⁸ It also proposed a moratorium on the anti-circumvention rules so that a study could be conducted about the potential impact of anti-circumvention rules on fair use, the public domain, and other noninfringing uses of copyrighted works.¹²⁹

More clearly than the Judiciary Committees in either branch of Congress, the Commerce Committee recognized the unprecedented nature of the access right that was implicit in the act-of-circumvention provision of section 1201. "If left unqualified," said Congressman Bliley, "this new right ... could well prove to be the legal foundation for a society in which information becomes available only on a 'pay-per-use' basis."¹³⁰ To ensure this would not occur, the legislation was amended to enable librarians and educators to make a showing that the anti-circumvention provision was interfering with noninfringing uses of copyrighted materials and to seek an exemption from the ban.¹³¹ Insofar as such a showing could be made, the Commerce Committee thought that affected classes of works or of users should be exempt from section 1201(a)(1)(A). Congressman Bliley pointed out that "[c]opyright law is not just about protecting information. It's just as much about affording reasonable access to it as a means of

125. See *Commerce Hearing*, *supra* note 44, at 1-3 (statement of Rep. Tauzin, Subcomm. Chairman) (explaining the Commerce Committee's reasons for reviewing the WIPO treaty implementation legislation).

126. See *Commerce Panel Clears Digital Copyright Bill With Further Concessions on Fair Use*, 56 BNA PAT., TRADEMARK & COPYRIGHT J. 326 (1998).

127. This eventually was codified in the DMCA. See 17 U.S.C.A. § 1201(g) (West Supp. 1999).

128. These were also eventually codified in the DMCA. See *id.* § 1201(h), (i).

129. See *id.* § 1201(a)(1)(B). See also *infra* notes 205-206 and accompanying text for discussion of this provision.

130. 144 CONG. REC. H7094 (daily ed. Aug. 4, 1998) (statement of Rep. Bliley).

131. See 17 U.S.C.A. § 1201(a)(1)(B)-(D). See *infra* notes 203-210 and accompanying text.

keeping our democracy healthy....”¹³² The Commerce Committee review of the legislation also led to inclusion of a provision indicating that nothing in section 1201 “shall enlarge or diminish any rights of free speech or of the press for activities using consumer electronics, telecommunications, or computing products.”¹³³ This provision recognizes the potential impact of the anti-circumvention rule on free speech and free press interests.

During the final negotiations leading up to passage of the DMCA, several of the exceptions were refined.¹³⁴ In addition, the computer security research community finally persuaded legislators to add another exception to enable circumvention of technical protection systems necessary for legitimate testing of the security of computer systems.¹³⁵

B. Circumvention for Other Legitimate Reasons Should Be Privileged

While the final version of the DMCA anti-circumvention provision responded to several significant concerns of the digital economy sector, it did so mainly by adopting specific exceptions. There are, however, many other legitimate reasons for circumventing technical protection systems that are not, strictly speaking, covered by the exceptions in the final bill. Five examples demonstrate that section 1201 should have an “or other legitimate purposes” exception to section 1201(a)(1).

Suppose, for example, that a copyright owner had reason to believe that an encrypted work contained an infringing version of one of its works. The only way to find out whether the copyright owner’s suspicion is valid may be to circumvent the technical protection system to get access to the encrypted material. Even if its suspicions proved correct, the copyright owner would have violated section 1201(a)(1)(A) in the course of discovering this. There is no exception in section 1201 to protect this kind of decryption activity.

Or suppose that a content producer had licensed certain software that was essential to the development of its product (e.g., editing software used in the process of making motion pictures). In the course of a dispute about the performance quality of this software, the content producer might with-

132. 144 CONG. REC. H7094 (daily ed. Aug. 4, 1998) (statement of Rep. Bliley).

133. 17 U.S.C.A. § 1201(c)(4).

134. Compare H.R. 2281, 105th Cong. (1998) (as passed on Aug. 4, 1998), with Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2860 (1998).

135. See 17 U.S.C.A. § 1201(j). This too had been the subject of testimony before the House Commerce Committee. See *Commerce Hearing*, *supra* note 44, at 27-30 (statement of Jonathan Callas).

hold payment of a royalty as a way of communicating its displeasure with the licensor's maintenance of the software. The software's licensor might then respond by activating a technical "self-help" system embedded in the software to stop the software from operating.¹³⁶ To deal with this development, the licensee might well attempt to circumvent the self-help feature now blocking access to the software because the licensee needed to use the software to finish its movie and because it regarded itself as having a legitimate claim of licensor breach to justify holding back the royalty.¹³⁷ However legitimate the claim or this activity, there is no exception to the anti-circumvention rule to protect the licensee in this situation.

Two further examples will illustrate the narrowness of certain existing privileges in the DMCA. Suppose, for example, that a firm circumvented a technical protection system to stop software it had licensed from monitoring certain uses of the software in ways not contemplated in the license agreement and which the licensee regarded as unwarranted and detrimental to its interests. Although there is a "personal privacy" exception in the DMCA,¹³⁸ there is no general exception for circumventing to protect other confidentiality interests. Or suppose that a firm was considering making a multi-million dollar acquisition of a computer system whose producer asserted was highly secure. If this firm wished to test the veracity of the producer's assertions, without getting the producer's permission or over the

136. Software developers can embed specialized disabling subprograms in licensed software. These may cause the software to cease operation unless a new code has been made available to the licensee by the licensor. They can also be invoked via a network connection to the licensor's site or by a remote act by the licensor. For a discussion of public policy issues raised by technical self-help systems, see Pamela Samuelson, *Embedding Technical Self-Help in Licensed Software*, 40 COMM. ACM 13 (1997).

137. A model law to regulate licensing of computer information has proposed to validate, as a matter of contract law, a licensor's use of technical self-help systems as long as certain procedural steps are taken to protect licensee interests. See U.C.C. § 2B-716 (Feb. 1999 Draft). See also Memorandum from Susan H. Nycum to Uniform Commercial Code Article 2B Reporter and Drafting Committee regarding Licensor Self-Help Revision of Proposed UCC 2B, at 1 (Jan. 27, 1997) available at <<http://www.2bguide.com/docs/nycshelp.html>> (expressing objections to proposed validation of technical self-help features in licensed software, speaking of them as a "trap for the unwary—in the extreme"); Memorandum from Michele Kane on behalf of Walt Disney Co. to Prof. Raymond T. Nimmer, Reporter for Article 2B, at 3 (Jan. 27, 1997), available at <<http://www.2bguide.com/docs/mkane.html>> (strenuously objecting to Article 2B's endorsement of technical self-help provisions in model licensing law as "unnecessary and unfair").

138. See 17 U.S.C.A. § 1201(i). For a discussion of the concerns leading to adoption of this exception, see *Commerce Hearing*, *supra* note 44, at 12-18 (statement of Marc Rotenberg, Director, Electronic Privacy Info. Ctr.).

producer's objection, it would seem to violate section 1201. Although there is a computer security testing exception in the Act, it only applies if one is already the owner or operator of the computer system being tested.¹³⁹ It should be noted here that many security flaws discovered in widely deployed systems have been found by researchers who tested the system without permission of either the owner or manufacturer of such systems.¹⁴⁰ These activities too are not covered by the computer security exception provided for in the DMCA.

Finally, because the DMCA recognizes that the anti-circumvention rules may have an impact on free speech and free press concerns,¹⁴¹ it may be worth considering an example of this sort. Suppose that an employee of a major chemical company gave a reporter a disk containing a digital copy of a report and several photographs pertaining to a major chemical spill that the company was trying to cover up. If information on the disk was technically protected and the employee was not authorized by the company to provide the information to the reporter, it would appear that the reporter would violate section 1201(a)(1) if he circumvented the technical protection system to get access to this information, even if consideration of free press and free speech interests might suggest that such a circumvention was justifiable.

One response to these examples might be to assert that copyright owners will generally not sue when these or other legitimate circumvention activities occur. However, in some of the examples given above, the technical protector might well have incentives to sue the circumventor.¹⁴² Given that there are serious criminal penalties for willfully violating section 1201,¹⁴³ the overbreadth of this provision and the narrowness of existing exceptions will put many legitimate circumventors at unnecessary risk. If such suits are brought, courts may, of course, and probably will, find other ways to reach just results. They might, for example, decide that the "other defenses" provision of the anti-circumvention rule legitimized the circumvention,¹⁴⁴ that some instances were within the spirit, even if not the letter, of an existing privilege, or that there was insufficient harm

139. See 17 U.S.C.A. § 1201(j).

140. See, e.g., John Markoff, *Software Security Flaw Puts Shoppers on Internet at Risk*, N.Y. TIMES, Sept. 19, 1995, at A1.

141. See 17 U.S.C.A. § 1201(c)(4).

142. See *supra* note 136 (licensor whose self-help feature might be defeated by a licensee).

143. See 17 U.S.C.A. § 1204.

144. See *id.* § 1201(c)(1).

to the legitimate interests of the person challenging the circumvention activity to justify imposing liability.¹⁴⁵ However, there should be a general purpose “or other legitimate purposes” provision in section 1201 so that courts will not have to thrash to reach appropriate results. This would add flexibility, adaptability, and fairness to the law. In many other parts of copyright law—with the fair use doctrine, for example, or the distinction between ideas and expressions—Congress has trusted the common law process to distinguish between legitimate and illegitimate activities. It could (and should) have done so with respect to circumvention legislation as well.

It would have been especially appropriate to adopt a general purpose “other legitimate purpose” provision because the anti-circumvention ban is an unprecedented provision for copyright law as to a significant new technology issue with which neither Congress nor the courts have much experience.¹⁴⁶ The lack of a general purpose exception is particularly troubling in view of the harsh criminal and civil provisions in the statute, which may have a chilling effect on legitimate activities, including those affecting free speech. It could also put at risk some legitimate activities in the digital economy that will impede the growth of e-commerce, as will become more apparent in the next section.

VI. THE ANTI-DEVICE PROVISIONS SHOULD BE NARROWED BY LEGISLATIVE AMENDMENT OR JUDICIAL INTERPRETATION

The text of the DMCA and its legislative history clearly demonstrate that Congress intended to ensure that users would continue to enjoy a wide range of noninfringing uses of copyrighted works, even if copyright owners used technical protection systems to impede them. This is evident in the DMCA’s recognition that circumventions for fair use, free speech, and

145. Section 1203(a) requires that a person be “injured by a violation of section 1201” in order to bring a suit to challenge a violation of this provision. *Id.* § 1203(a).

146. Professor Julie Cohen, in commenting on the structure of section 1201, observed that this provision is almost European in its construction. Typically, European legislators formulate laws as though all contingencies can be foreseen and the rule can be established for all time. Europeans typically provide a broad rule and only limited exceptions to the rule. American laws more typically have some openness that allow the laws to adapt to new circumstances. This may provide American law with needed flexibility in times of rapid technological change. Yet, section 1201 deviates from this general American approach. Conversation with Julie E. Cohen (Jan. 1999).

free press purposes should be lawful.¹⁴⁷ It is also apparent in the provision enabling the Librarian of Congress to exempt certain classes of users or works from the general anti-circumvention rule when necessary to preserve socially valued noninfringing uses.¹⁴⁸ In addition, it explains why Congress adopted some exceptions to the act-of-circumvention ban, notably, the interoperability privilege.¹⁴⁹ As the last part has shown, if Congress had not been blinded by the politics of the day, it would likely have recognized other legitimate reasons to engage in acts of circumvention.

If Congress intended for circumvention of technical protection systems to be legal when done for legitimate purposes, it might seem obvious that Congress should be understood to have intended to enable users to effectuate the circumvention privileges it recognized.¹⁵⁰ Although it will not always be necessary for a legitimate circumventor to make or use a circumvention technology to accomplish a privileged circumvention (e.g., enciphered text might be decoded by purely mental activity), most often this will be necessary.¹⁵¹ The deepest puzzle of section 1201 is whether Congress implicitly intended to allow the development and/or distribution of technologies necessary to accomplish legitimate circumvention activities, or whether, in essence, it created a number of meaningless privileges.

Seemingly relevant to addressing this question are some curious features of section 1201 that close study of this complex provision reveals. First, several exceptions to the anti-circumvention rule specifically authorize the creation of tools necessary to achieving a legitimate circumvention activity (e.g., the encryption research and interoperability privi-

147. See 17 U.S.C.A. § 1201(c)(1), (c)(4), discussed *supra* notes 99, 107, 113-116 and accompanying text. This same subsection indicates that it also does not intend to enlarge or diminish vicarious or contributory copyright infringement. See *id.* § 1201(c)(2).

148. See *id.* § 1201(a)(1)(B)-(D).

149. See *id.* § 1201(f). This exception preserves the fair use privilege recognized in *Sega Enterprises, Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992), that permits the intermediate copying of computer programs when necessary to obtain information in order to achieve interoperability among independently developed computer programs.

150. See Benkler, *supra* note 24, at 416 ("If the act of circumvention were privileged to users, particularly if it were privileged as a matter of free speech, it would be difficult to sustain a prohibition on manufacture and sale of the products necessary to enable users to engage in circumvention.").

151. See, e.g., James R. Davis, *On Self-Enforcing Contracts, the Right to Hack, and Willfully Ignorant Agents*, 13 BERKELEY TECH. L.J. 1145, 1147 (1998) (questioning whether a "right to hack" for fair use would be meaningful, given that most users would be unable to overcome technical protection systems without tools designed for that purpose).

leges),¹⁵² while several others (e.g., the law enforcement privilege and the privacy privilege) do not.¹⁵³ Secondly, while the interoperability privilege exempts necessary tools from both device provisions of section 1201,¹⁵⁴ the encryption and security research privileges exempt tools only from the access-device provision, not from the control-device provision. Yet, it would seem that encryption and security research would often require testing both of access and of control components of technical protection systems.¹⁵⁵ Thirdly, section 1201 contains no provision enabling the development or distribution of circumvention tools to enable fair use or other privileged uses in terrain which section 1201(a)(1)(A) doesn't reach (i.e., making fair uses of lawfully acquired copies). If Congress intended to recognize a right to "hack" a technical protection system to make fair uses, this right could be undermined if it could not be exercised without developing a tool to bypass the technical protection system or otherwise getting access to such a tool.¹⁵⁶ Under some interpretations of section 1201(b)(1), development or distribution of such a tool would be unlawful.

Consider, for example, the Xerox PARC researcher who circumvented a movie's technical protection system in order to make a fair use clip for the Washington Redskins' litigation.¹⁵⁷ It was necessary for him to develop a tool to enable him to bypass the technical protection system to make the clip. Suppose that the motion picture copyright owner found out about the circumvention and decided to make an example of this researcher, suing him for statutory damages for violating section 1201(b)(1).¹⁵⁸ On a strict interpretation of this subsection, the researcher might seem to be in trouble. The tool was, after all, "primarily designed ... for the purpose of circumventing protection afforded by a technological

152. See 17 U.S.C.A. § 1201(f)(2), (g)(4).

153. See *id.* § 1201(e), (i). There is, however, a better textual argument for inferring a tool-making privilege for law enforcement activities than for inferring tool-making authority to enable privacy protection. Section 1201(i) limits the application of section 1201(a)(1)(A), whereas § 1201(e) indicates that "this section does not prohibit any lawfully authorized investigative ... activity" of a government agent.

154. See *id.* § 1201(f)(2).

155. See *id.* § 1201(g)(4), (j)(4).

156. See Cohen, *supra* note 63, at 174-78 (discussing lawful tampering with technical protection systems and its implications for the availability of tools to accomplish this).

157. See *supra* note 114-116 and accompanying text.

158. See 17 U.S.C.A. § 1203(c)(3). This researcher would likely be spared from criminal liability for violation of § 1201(b) because he was serving as a *pro bono publico* expert witness in this case. Section 1204(a) requires that a violation of § 1201 not only be willful, but done for commercial advantage or private financial gain for criminal liability to be imposed. See *id.* § 1204(a).

measure that effectively protects a right of the copyright owner under this title in a work or a portion thereof.”¹⁵⁹ However, one can easily imagine a court deciding that the researcher’s code did not run afoul of section 1201(b)(1). The code might be viewed as a special purpose tool made for the limited purpose of effectuating fair use rights. In view of its lack of commercial significance and the absence of deleterious effects of the sort that the anti-device provisions were intended to reach,¹⁶⁰ a court might decide that this code should not be held to violate this law.¹⁶¹

Would the result be different if the researcher asked a co-worker or a friend to develop the tool instead of doing it himself? Or would the result be different if the researcher shared this tool with a co-worker who needed to make a fair use circumvention of a different movie? Even though he might be “provid[ing]” this technology to another person, perhaps he would escape liability because he was not “traffic[king]” in this technology or “offer[ing it] for sale” which are the principal activities Congress meant to curb by enacting this part of DMCA.¹⁶² However, it is fair to observe that courts would have to read some limiting language into section 1201(b)(1) to decide that the researcher would not be liable in all three situations.

An undoubtedly closer question is what courts would do about a technology distributed in the mass-market for purposes of enabling privileged circumventions. Consider, for example, how the 1985 *Vault v. Quaid*¹⁶³ case would fare under the DMCA anti-device provisions. Vault sued Quaid for contributory copyright infringement based on Quaid’s development and sale of a program called Ramkey. Quaid’s customers could use Ramkey to defeat Vault’s Prolok copy-protection software (which Vault sold to other software developers to protect their own software from unauthorized copying). By spoofing Vault’s copy-protect system,¹⁶⁴ Quaid’s customers could make unauthorized copies of the third-party software protected by Vault’s program.¹⁶⁵ Quaid successfully defended against the

159. *Id.* § 1201(b)(1).

160. See House Manager’s Report, *supra* note 63, at 9-13.

161. Alternatively, the court could find only a technical or de minimis violation of the statute in this instance.

162. 17 U.S.C.A. § 1201(b)(1).

163. 775 F.2d 638 (5th Cir. 1985).

164. In essence, this and other “spoofing” software generally operate by emitting a signal which will be interpreted by the other firm’s copy-protection software (or conceivably hardware) as an indication that the system is operating effectively.

165. Vault also claimed direct copyright infringement, trade secret misappropriation, and breach of contract. See *Vault*, 847 F.2d at 257-58.

contributory infringement claim by showing that Ramkey had a substantial noninfringing use, namely, to enable users to effectuate their rights under copyright law to make backup copies.¹⁶⁶

Quaid would probably not run afoul of the access-device provision of section 1201(a)(2).¹⁶⁷ However, less clear is whether it could successfully defend against a section 1201(b)(1) claim. Suppose that Quaid's president testified that his primary purpose in designing and producing Ramkey was to enable his customers to do legitimate backup copying. Suppose further that the marketing literature for Ramkey emphasized this purpose of the program and even warned potential customers not to use Ramkey to make infringing copies. If a court considered this evidence credible, it would probably save Quaid from criminal prosecution for violating the second anti-device norm, because it would show a lack of wrongful intent. But would it save Quaid from civil liability?¹⁶⁸

To answer that question, courts would have to grapple with a seeming inconsistency in the statute. On the one hand, the DMCA seems to outlaw technologies if their primary purpose is to circumvent a technical protection measure that effectively protects a right of a copyright owner to con-

166. *See id.* at 262 (relying on the Supreme Court's decision in *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417 (1984), which rejected a claim that Sony had contributorily infringed Universal's movie copyrights by selling Betamax machines which enabled home copying of these movies off the broadcast television because of noninfringing uses of the Betamax machine).

167. Quaid could probably argue that Ramkey was primarily designed to enable bypassing of the Prolok system for lawfully acquired copies of protected programs. This would seem to make § 1201(a)(2) inapplicable to the *Vault v. Quaid*-like controversies.

168. An interesting question is who could sue Quaid under § 1201(b)(1). The Clinton Administration's Green Paper on Intellectual Property and the National Information Infrastructure suggested that the maker of a protective technology, such as Vault, would not have standing to challenge the maker of circumvention technologies. *See* U.S. GOV'T WORKING GROUP ON INTELLECTUAL PROPERTY, GREEN PAPER ON INTELLECTUAL PROPERTY AND THE NATIONAL INFORMATION INFRASTRUCTURE 130 (1994). Copyright owners who used technical protection systems to protect their works would seem to have standing to initiate the suit. This could mean that a firm such as Quaid would thus be faced, not just with one lawsuit, but potentially thousands to defend. As will be discussed further, *see infra* note 194 and accompanying text, in none of these lawsuits would the plaintiff have to demonstrate that any underlying act of infringement actually took place. The White Paper was silent on the issue of standing. Nor is the issue expressly dealt with in the DMCA. Proposals by representatives of Macrovision Corp., which makes technical protection systems, to change 17 U.S.C.A. § 1203(a) to facilitate its ability to obtain standing in such a suit were not heeded by Congress. *See Judiciary Hearing, supra* note 17, at 271-77 (statement of Mark S. Belinsky, Vice President, Copy Protection Group, Macrovision Corp.).

trol its work (in this case, a right to control illegal copying).¹⁶⁹ On the other hand, the DMCA recognizes that fair use-like circumventions should be lawful.¹⁷⁰ Backup copying is a specially privileged activity in the copyright statute.¹⁷¹ Because the copyright owner doesn't have a statutory right to control backup copying, perhaps a spoofing technology intended to enable backup copying should be outside the statute. It is important to understand that circumvention for backup copying purposes generally cannot occur without access to such a technology.

So if most lawful users of Prolok-protected software lack the skills to write a Ramkey-equivalent, perhaps it should be lawful to make and distribute a technology to effectuate the backup copy privilege. It is unclear whether Congress intended for the technologically savvy who could "do it themselves" to be the only ones who could engage in privileged acts of circumvention. Yet, as the example of the Xerox researcher illustrates, even the technically sophisticated will often need to develop a tool to accomplish a privileged circumvention; this would seem to put them at risk under a strict reading of section 1201(b)(1).¹⁷²

Potentially relevant to whether the distribution of a tool like Ramkey is lawful is section 1201 (c)(2), which states that nothing in section 1201 "shall enlarge or diminish vicarious or contributory liability for copyright infringement in connection with any technology, product, service, device, component, or part thereof."¹⁷³ If what this subsection purports is true, perhaps the result in *Vault v. Quaid* would be the same after DMCA as before. One can imagine some courts deciding to construe section 1201(b)(1) narrowly so that the honest maker of a Ramkey-equivalent for purposes of enabling backup copying would be able to do so. But they are certainly not constrained to do so.

Moreover, the major copyright industries that supported a broad ban on circumvention technologies would assert that courts should not construe the DMCA so narrowly. They would likely consider *Quaid's* argu-

169. See 17 U.S.C.A. § 1201(a)(2), (b)(1).

170. See *id.* § 1201(c)(1), discussed *supra* notes 99, 107, 113-116, 147 and accompanying text.

171. See 17 U.S.C. § 117 (1994).

172. Even they, of course, may have to manufacture a technology or provide a service to make backup copies, in apparent violation of section 1201's anti-device rules. See Benkler, *supra* note 24, at 416.

173. 17 U.S.C.A. § 1201(c)(3). Recall that the main claim made by *Vault* against *Quaid* was a contributory infringement claim, and it was unsuccessful. See *supra* note 163-166 and accompanying text.

ment that Ramkey was primarily designed and produced to enable lawful backup copying as a ruse. Moreover, they would likely point out that Ramkey doesn't just enable lawful backup copying; it enables illegal copying as well. They would regard the danger that Ramkey would be used for illegal purposes—regardless of Quaid's intent—as so substantial as to justify banning this technology. The DMCA's anti-device provisions were broadly drafted, they would argue, to address this very danger.¹⁷⁴ They would also consider it an unnecessary burden for copyright owners to have to prove that the primary use of a technology like Ramkey was to engage in infringement.¹⁷⁵ This would be difficult to do, especially for a technology that was about to be introduced into the market. When the dangers of infringement are high, they would argue, the technology ought to be deemed illegal if its purpose is to circumvent a technical protection system copyright owners are using to protect rights granted to them by copyright law.¹⁷⁶ According to this view, Ramkey is illegal under the DMCA. The major copyright industry supporters of the broad anti-device provisions of the DMCA would probably like nothing better than to make Congress' apparent preservation of noninfringing uses into a meaningless promise.

Different judges might reach different conclusions on a Ramkey-like case, but consider how they might deal with another plausible "spoofing" technology. Intel has recently developed a line of semiconductor chips with a built-in identification system for each processor.¹⁷⁷ Privacy advocates have raised concerns about the threat that processor identification systems pose for personal privacy on the Internet.¹⁷⁸ In response to these

174. See *Judiciary Hearing*, *supra* note 17, at 57 (statements of Hon. Bruce A. Lehman, Commissioner of Patents and Trademarks, Patent and Trademark Office).

175. See *Commerce Hearing*, *supra* note 44, at 54-58 (prepared statement of Steven J Metalitz on behalf of the Motion Picture Ass'n of America) (objecting to proposals that would require copyright owners to prove that circumvention or circumvention devices would cause infringement).

176. There is no "authority of law" exception in the DMCA's anti-device provisions, as there was in the White Paper's original proposal for an anti-device regulation. See White Paper, *supra* note 15, app. 1 at 6. How, if at all, this might affect the scope of the DMCA's anti-device provisions remains to be seen.

177. See Peter H. Lewis, *Whoosh! The Next Pentium Chip Is On Its Way*, N.Y. TIMES ON THE WEB (Jan. 14, 1999) <<http://www.nytimes.com/library/tech/99/01/circuits/articles/12pete.html>>.

178. See Jeri Clausing, *Privacy Groups Seek Recall of Intel Chip*, N.Y. TIMES ON THE WEB (Jan. 29, 1999) <<http://www.nytimes.com/library/tech/99/01/cyber/articles/29privacy.html>>. Although the threat the Intel processor ID poses for privacy has gotten the most attention in the press, the potential for the Intel processor ID to be used to pre-

concerns, Intel announced its intent to ship these chips with the processor identity function “off.”¹⁷⁹ Suppose, however, that Microsoft develops Windows 2000 as a “trusted system” technology¹⁸⁰ to run on this line of Intel chips and that it requires that licensees of Windows 2000 agree to keep the Intel identification system on at all times.¹⁸¹ Having the identifier on, Microsoft might well contend, is a critical component to the effectiveness of its trusted system technology. Suppose further that Windows 2000 will not install until the Intel identifier is on, and that the installation software, after a user clicks “I agree” to the conditions of the license, will actually turn the identifier on if necessary.¹⁸² If a privacy advocacy group developed and distributed software to spoof Windows into thinking the Intel identifier was on when it was not in order to protect user privacy, or if the group posted information about how users could turn the identifier off even when using Windows 2000, would it be violating section 1201(b)(1)?¹⁸³

vent “piracy” of software has also been recognized. See Peter Wayner, *Debate on Intel Chip Misses Piracy Issue*, N.Y. TIMES ON THE WEB (Jan. 30, 1999) <<http://www.nytimes.com/library/tech/99/01/cyber/articles/30chip.html>>.

179. See Jeri Clausing, *Intel Alters Plan Said to Undermine PC Users' Privacy*, N.Y. Times, Jan. 26, 1999, at A1.

180. “Trusted system” is a term used to describe a computer and software system constructed to make it impossible (or at least very difficult) to make unauthorized copies or uses of legally protected works. See Mark Stefik, *Shifting the Possible: How Trusted Systems and Digital Property Rights Challenge Us To Rethink Digital Publishing*, 12 BERKELEY TECH. L.J. 137 (1997).

181. This is no mere conjecture. Microsoft is reportedly intending to deploy trusted system software with the next version of Windows. See Jason Chicola et al., *Digital Rights Architectures for Intellectual Property Protection 99* (1998), paper prepared for *Ethics and Law on the Electronic Frontier*, Massachusetts Institute of Technology, available at <<http://swissnet.ai.mit.edu/6805/student-papers/fall98-papers/trusted-systems/trustsys.doc>> (MS Word document). This is especially worrisome since Microsoft has a monopoly position in the market for operating systems software, making it largely immune from competitive pressures that might limit its ability to impose trusted system technology on the market.

182. Another important policy initiative affecting the enforceability of mass-market licenses of this sort is proposed Article 2B of the Uniform Commercial Code. See generally Symposium, *Intellectual Property and Contract Law in the Information Age: The Impact of Article 2B of the Uniform Commercial Code on the Future of Transactions in Information and Electronic Commerce*, 13 BERKELEY TECH. L.J. 809 (1998); Symposium, *Intellectual Property and Contract Law for the Information Age: The Impact of Article 2B of the Uniform Commercial Code on the Future of Information and Commerce*, 87 CALIF. L. REV. 1 (1999).

183. If the Pentium III chip ID is designed to allow for copyright protection, as Intel claims it is, it might be a technology which effectively controls access to copyrighted

Under a very strict interpretation of section 1201(b)(1), either act might be viewed as illegal.¹⁸⁴ It is, however, difficult to believe that most judges would find providing either software or information to enable circumvention of this component of a technical protection system to fall within the DMCA anti-device rules. The DMCA, judges might point out, authorizes circumvention in order to protect personal privacy.¹⁸⁵ While this provision doesn't specifically authorize the development or use of circumvention technologies to accomplish this legitimate act, judges might conclude that Congress must have intended for people to be able to develop or use technology to accomplish the privileged privacy act, or that the Intel identifier was not a component of an effective technical measure. To avert an injustice, judges would likely find an ambiguity in the statute or read in appropriate limiting language. This is clearly not the kind of "black box" circumvention device that Congress had in mind when adopting DMCA.¹⁸⁶ To hold otherwise would, in effect, allow Microsoft to employ the anti-circumvention provisions of DMCA to impose trusted system technology on the public.

It is, of course, an irony that so much of Congressional debate on section 1201 focused on refining the act-of-circumvention provision given that the anti-device provisions are, as a practical matter, by far the more important rules in this section.¹⁸⁷ Those who have followed the Clinton Administration's digital copyright policy over the last five years should realize that the anti-device provisions were what Administration officials and major copyright industry allies really cared about. The legislation proposed in the Administration's 1995 White Paper did not include any provision about circumvention of technical protection measures as such.¹⁸⁸ It sought only to outlaw technologies whose "primary purpose or effect" was

works under § 1201. If so, it would seem that a hardware device which disables the Processor Serial Number could be subject to the anti-device provisions. Take, for example, IBM's new hardware disablement feature: "IBM plans to go the extra step and disable the processor ID feature at the BIOS (or hardware) level in our Pentium III client systems," Letter from Christopher G. Caine on behalf of IBM Corp. to Jerry Berman, Executive Director of the Center for Democracy and Technology (Jan. 24, 1999), *available at* <<http://www.cdt.org/privacy/ibmletter.shtml>>.

184. Posting information on the website might be seen as providing a service to the circumventors.

185. See 17 U.S.C.A. § 1201(i) (West Supp. 1999). This provision is extremely complicated and would seem to be very narrow. It is not clear it would apply to the Microsoft example.

186. See *supra* note 93 and accompanying text and *infra* note 231.

187. See Benkler, *supra* note 24, at 416.

188. See White Paper, *supra* note 15, at 230-36.

to enable the circumvention of technical protection measures.¹⁸⁹ Was this lack of attention to circumvention an oversight? Or did the Administration believe that it would be difficult to detect individual acts of circumvention, and as long as such acts were done on an isolated, individual basis (due to the unavailability of circumvention devices), the danger to copyright owners would be small? It is difficult to discern why circumvention as such escaped attention until mid-1997 when the WIPO treaty implementation legislation was first introduced in Congress.¹⁹⁰ Far easier to discern has been the Administration's goal of stopping the manufacture and distribution of technologies with circumvention-enabling uses, either by commercial firms or by technically savvy Robin Hoods.¹⁹¹

Eventually someone in the Administration must have realized that it was a bit strange to be proposing to make illegal the manufacture and distribution of technologies whose ordinary uses were not themselves illegal. To justify a broad ban on circumvention technologies, a broad ban on the act of circumvention seemed to be needed. This explains why the Administration and its allies were so insistent that section 1201(a)(1) be structured to broadly ban acts of circumvention. It also explains why the Administration sought to limit the proliferation of exceptions to the anti-circumvention ban, and why such exceptions as were added to the statute were very narrow. The broader the acknowledged range of legitimate circumventions, the narrower should be an appropriately crafted regulation of circumvention technologies. The Administration may have hoped that in all the hoopla about crafting exceptions to section 1201(a), Congress would not notice that its seeming recognition of the legitimacy of circumventions for noninfringing purposes in section 1201(c)(1) might effectively be nullified by section 1201(b)(1), which arguably broadly bans technologies necessary to accomplish such circumventions.

When testifying before Congress, proponents of the Administration's anti-device rules repeatedly emphasized that the legislation was needed to stop deliberate and systematic piracy by "black box" providers.¹⁹² Yet, the

189. See *id.*, app. 1 at 6.

190. See *supra* note 86.

191. Professor Benkler likens this strategy to banning VCRs in order to stop home taping. See Benkler, *supra* note 24, at 416. Speaking of VCRs, little noticed in DMCA were its provisions requiring consumer electronics companies to build specific anti-copying technologies into future VCRs. See 17 U.S.C.A. § 1201(k) (West Supp. 1999).

192. See *Judiciary Hearing*, *supra* note 17, at 212-16 (statement of Gail Markels, General Counsel and Senior Vice President, Interactive Digital Software Ass'n) (relying on example of circumvention device with no legitimate purpose that had been used to pirate games); *id.* at 273-77 (prepared statement of Mark Belinsky on behalf of Macrovi-

anti-device provisions adopted by Congress are far broader than this, providing a basis to challenge an unacceptably wide range of technologies that have circumvention-enabling uses. This creates a potential for “strike suits” by nervous or opportunistic copyright owners who might challenge (or threaten to challenge) the deployment of a new information technology tool whose capabilities may include circumvention of some technical protection system. No doubt some expert can be found to say that deployment of a particular technology in the market would meet one of the three conditions in the anti-device provisions, giving plausibility to the suit. Weak as such testimony might be, it may be enough to extract a settlement sum from the information technology firm.¹⁹³

The potential for strike suits becomes stronger if one realizes that it is not necessary (or arguably even relevant) to litigation under the anti-device provisions of DMCA whether any act of underlying infringement (e.g., illegal copying of a protected work) has ever taken place. The mere potentiality for infringement will suffice to confer rich rewards on a successful plaintiff. Consider, for example, a recent lawsuit brought by the maker of a proprietary game console against the maker of emulation software that permits games initially developed for the proprietary console to be played on iMac computers.¹⁹⁴ Relying on the DMCA anti-device provision, the plaintiff is seeking up to \$25,000 per unit sold in damages because the emulation software allegedly bypasses an anti-copying feature in

sion Corp.) (emphasizing the need to outlaw pirate devices). *See also NII Copyright Protection Act of 1995 (Part II): Hearings on H.R. 2441 Before the Subcomm. on Courts and Intellectual Property of the House Comm. on the Judiciary*, 104th Cong. 23 (1996) (prepared statement of Jack Valenti, President and CEO, Motion Picture Ass'n of America) (“But all security measures, no matter how sophisticated, can be circumvented by clever hackers. Therefore, the law must provide clear and effective sanctions against those who would violate the security of the NII. This requires more than mere civil remedies. Criminal sanctions are essential. Too many NII bandits, some operating totally in the underground economy, will scoff at the threat of civil damages, which many regard as simply a cost of doing business. There must be criminal penalties attached to deliberate, systematic acts of circumvention if such acts are to be seriously lessened.”).

193. Some commentators even perceive the anti-device rules of § 1201 as threatening the distribution of many widely used editing and related software tools. *See Peter Wainer, The Copyright Boomerang*, SALON MAGAZINE (Nov. 20, 1998) <<http://www.salonmagazine.com/21st/feature/1998/11/20feature.html>> (considering whether “cutting and pasting” will be rendered unlawful).

194. *See* Complaint, Sony Computer Entertainment, Inc. v. Connectix Corp., Civ., No. 99-0390 (N.D. Cal., filed Jan. 27, 1999) [hereinafter Sony Complaint]. For a discussion of this lawsuit, see Band & Isshiki, *supra* note 65.

the games.¹⁹⁵ The plaintiff did not allege and need not prove any actual illicit copying by users of the defendant's emulation software.

The anti-device provisions of section 1201 are not predictable, minimalist, consistent, or simple, as the Framework principles suggest that they should be. Due to inconsistencies in the statute, it is unclear whether section 1201's anti-device provisions would be interpreted to allow the development and distribution of technologies to enable legitimate uses. Boiled down to its essence, this presents the question of whether Congress should be understood to have made an empty promise of fair use and other privileged circumvention. Unless the anti-device provisions of the DMCA are modified, either by narrow judicial interpretation or by legislative amendments,¹⁹⁶ they are likely to have harmful effects on competition and innovation in the high technology sector. This is not good news for the digital economy.

VII. POLICYMAKERS SHOULD PERIODICALLY REVIEW BOTH THE ACT AND DEVICE PROVISIONS

The Clinton Administration did not recommend or support inclusion of any provision in the WIPO treaty implementation legislation to assess the impact of the DMCA's anti-circumvention norms.¹⁹⁷ This might seem surprising in view of the breadth of these norms, their unprecedented character, and their potential impact on both information technology markets and on public access to information. Even if the Administration had initially been unaware of these potential negative impacts, it could not have failed to become aware of them during the legislative process.¹⁹⁸ The Administration was surely aware that the case for the act-of-circumvention and anti-device norms was long on rhetoric and short on actual evidence of

195. See Sony Complaint, *supra* note 194, at 7-8. This lawsuit is particularly disturbing because the software at issue was named "Best of Show" at Macworld Expo shortly before the lawsuit was filed. See *Best of Show*, MACWORLD ONLINE (visited Apr. 21, 1999) <<http://macworld.zdnet.com/expo/report/bestofshow.html>>.

196. A predictable, minimalist, consistent, and simple anti-device norm might outlaw the manufacture and distribution of technologies intended to facilitate copyright infringement or of technologies with limited legitimate uses.

197. See H.R. 2281, 105th Cong. (1997) (as originally introduced into Congress on July 29, 1997); *Judiciary Hearing*, *supra* note 17, at 34-42 (statement of Bruce Lehman) (endorsing legislation but not asking for a study provision).

198. See *Judiciary Hearing*, *supra* note 17, at 148-56 (statement of Robert L. Oakley); *id.* at 64-68 (statement of M.R.C. Greenwood).

harm to copyright owners.¹⁹⁹ Yet, the Administration did nothing to support post-legislative review of these norms.

This is in striking contrast to the periodic review process endorsed by the Administration as to another legislative initiative affecting digital economy markets, namely, the proposal to create a new form of legal protection for the contents of databases.²⁰⁰ Writing on behalf of the Administration concerning its reservations about a bill under consideration in the second session of the 105th Congress, Andrew Pincus, General Counsel to the Commerce Department, stated:

The Administration believes that, given our limited understanding of the future digital environment and the evolving markets for information, it would be desirable for the [database] bill to include a provision for an interagency review of the law's impact at periodic intervals following implementation of the law. This would be consistent with the laws and proposed laws in other emerging technologies where Congress has mandated examination of a new law's economic impact.²⁰¹

At least one of the database bills seemingly under consideration in the 106th Congress contains a study provision to assess the impact of the new law.²⁰² This conforms to the Administration's proposal and to Framework principles. Much the same comment might have been made about the anti-circumvention norms of the DMCA.

Even though the Administration did not support inclusion of study provisions in the DMCA, section 1201 actually does contain a study provision that will provide an opportunity to review some impacts of the anti-

199. One of the few concrete examples of a device claimed to have contributed to international piracy was that offered in *Judiciary Hearing*, *supra* note 17, at 213-216 (statement of Gail Markels) (discussing "Game Doctor" said to have been used to pirate game software in Hong Kong and Taiwan).

200. See Letter from Andrew Pincus, General Counsel of the U.S. Dep't of Commerce, to Sen. Patrick Leahy 3 (Aug. 4, 1998) (on file with author) [hereinafter Pincus Letter]. After the House passed the Collections of Information Antipiracy Act, H.R. 2652, 105th Cong. (1998), Mr. Pincus wrote to Senator Leahy to express the Administration's reservations about the wisdom of this bill and about its constitutionality. See Pincus Letter, *supra*, at 1.

201. Pincus Letter, *supra* note 200, at 3. The letter proposed that "such a study might be conducted under the auspices of the Secretary of Commerce in consultation with the Office of Science and Technology Policy and the Register of Copyrights." *Id.*

202. See 145 CONG. REC. S322 (daily ed. Jan. 19, 1999) (provision entitled "Report to Congress," from one of three potential database bills referred to by Sen. Hatch).

circumvention regulations.²⁰³ In response to the strong concerns expressed by librarians and educators about the potential negative impacts that broad anti-circumvention provisions might have on fair uses of copyrighted works and on access to information and to public domain works,²⁰⁴ the House Commerce Committee decided that there should be a two-year moratorium on enforcement of the act-of-circumvention provision.²⁰⁵ It also proposed a study to determine whether noninfringing uses were being adversely affected by technical protection systems. If so, the Commerce Committee's version of the bill would have waived application of the anti-circumvention norm as to the affected works or users.²⁰⁶

The Commerce Committee's insistence on the moratorium and an impact study proved surprisingly persuasive. Section 1201(a)(1)(A) provides that the general anti-circumvention ban will not take effect until two years after enactment of the legislation.²⁰⁷ Subsections (C) and (D) call upon the Librarian of Congress to conduct periodic studies to determine whether certain classes of users or works should be exempt from the ban because technical protection systems are impeding the ability to make noninfringing uses of copyrighted works.²⁰⁸ Subsection (B) goes on to provide the statutory basis for granting such an exemption to the classes of works or users determined by the Librarian to be adversely affected by the anti-circumvention norm.²⁰⁹ The DMCA calls for the Librarian's first study to be completed before the anti-circumvention moratorium ends.²¹⁰

203. See 17 U.S.C.A. § 1201(a)(1)(B)-(D) (West Supp. 1999). Section 1201 also contains a provision for studying the impact of the encryption research provision. *Id.* § 1201(f)(5).

204. See *supra* note 117 and accompanying text.

205. See *Commerce Panel Clears Digital Copyright Bill With Further Concessions On Fair Use*, 56 BNA PAT., TRADEMARK & COPYRIGHT J. 326, 326 (1998).

206. See *id.* As Professor Benkler has pointed out, this would not stop copyright owners from employing technical protection systems to inhibit noninfringing uses; it would only allow circumvention to obtain access. See Benkler, *supra* note 24, at 428.

207. 17 U.S.C.A. § 1201(a)(1)(A) (West Supp. 1999).

208. The first study is to be completed two years after the date of DMCA's enactment. See 17 U.S.C.A. § 1201(a)(1)(A) (West Supp. 1999). Follow-on studies are to be conducted every three years thereafter. See *id.* § 1201(a)(1)(C). Given how weak was the showing that gave rise to the DMCA's anti-device provisions, it would seem that the showing of interference with lawful uses ought not to be too rigorous. However, the House Manager's report on the legislation would seem to anticipate a relatively high standard of proof. See House Manager's Report, *supra* note 63, at 6-7.

209. See 17 U.S.C.A. § 1201(a)(1)(B) (West Supp. 1999). It appears that any moratorium resulting from such a determination will last for three years. *Id.* § 1201(a)(1)(D). The rulemaking procedure set forth in § 1201(a)(1)(B)-(D) may, however, be unconstitu-

The DMCA directs the Librarian of Congress to consider four factors in carrying out this study:

(i) the availability for use of copyrighted works, (ii) the availability for use of works for nonprofit archival, preservation, and educational purposes, (iii) the impact [of] the prohibition ... on criticism, comment, news reporting, teaching, scholarship, or research, [and] (iv) the effect of circumvention of technical measures on the market for or value of copyrighted works."²¹¹

The Librarian has authority to consider "such other factors as the Librarian considers appropriate."²¹² The DMCA is quite clear, however, that the Librarian's determinations cannot be asserted as a defense to an anti-device claim.²¹³ Although users would be entitled, after the Librarian's determination, to "hack" technical protection systems for any classes of works whose noninfringing uses had been inhibited, the no-defense-to-an-anti-device-claim subsection would appear to make such user self-help available only if one could accomplish the act without a device, once again raising the specter of Congress having created a meaningless privilege. As Professor Benkler has pointed out, the Librarian has no power to tell copyright owners to stop using technical protection systems that are impeding noninfringing uses.²¹⁴ Thus, it is quite possible that noninfringing uses will continue to be substantially impeded, notwithstanding the Librarian's determination and rulemaking concomitant to it. Surely, this should be the subject of further study.

While the study provisions in DMCA are surely better than nothing,²¹⁵ they fall far short of the periodic review and reporting process appropriate to the unprecedented nature of the anti-circumvention regulations.²¹⁶ To

tional because the Librarian of Congress is not an executive branch official. *See* Band & Isshiki, *supra* note 65, at 7.

210. *See* 17 U.S.C.A. § 1201(a)(1)(A) (West Supp. 1999).

211. *Id.* § 1201(a)(1)(C). Another subsection of the DMCA requires the Register of Copyrights and the Assistant Secretary for Communications and Information of the Commerce Department to study the impact of the encryption research exception. *See id.* § 1201(g)(5).

212. *Id.* § 1201(a)(1)(C).

213. *See id.* § 1201(a)(1)(E).

214. *See* Benkler, *supra* note 24, at 428.

215. The principal value of the study provisions may well lie in deterring some publishers from egregious uses of technical protection systems that would interfere with fair uses..

216. Among the factors likely to limit the effectiveness of the study system devised in the DMCA is the fact that the Librarian of Congress is apparently supposed to initiate

limit an assessment of the circumvention ban to a narrow range of possible effects would ignore the wider swath of harm the provision may do.²¹⁷ Besides, the device ban is the true heart of the anti-circumvention provisions of the DMCA. It is integrally interrelated with the circumvention activity ban.²¹⁸ To assess the act-of-circumvention ban without considering the device ban is to ignore the most significant technology-regulating provision in the DMCA. Unless construed narrowly, the anti-device provisions may do as much harm to competition and innovation in the information technology industry as the circumvention ban may do to noninfringing academic uses. One would have thought that Congress and the Administration would be concerned about these impacts given that these are the very industries whose tremendous growth the Commerce Department has been trumpeting to the world.²¹⁹ The Librarian of Congress should, therefore, decide that his studies can consider the impact of anti-device rules on the ability of certain classes of users or works to make noninfringing uses of protected works.²²⁰ The Librarian should also be entitled to make other observations about possible unintended side-effects of the anti-circumvention regulations that may be detrimental to the public interest.²²¹

It is especially important to have periodic reviews of the whole of the anti-circumvention provisions because they sweep so broadly that they may come to be used widely to deal with circumventions far beyond the copyright industry concerns that Congress contemplated. The low level of

studies of the impact of anti-circumvention rules "upon the recommendation of the Register of Copyrights." *Id.* § 1201(a)(1)(C). The Register, in turn, is supposed to consult with an official from the Department of Commerce before recommending a study. *See id.* It has been a long time since the Register of Copyrights or the Commerce Department have taken more than tepid steps to represent the interests of users of copyrighted works, particularly those from the educational and library sectors. Moreover, because none of the Librarian's findings last for more than a three year period, copyright industry lobbyists will have multiple opportunities to carve back or eliminate any user-friendly exceptions that the Librarian might have the temerity to recommend.

217. *See supra* note 136-140 and accompanying text for examples of legitimate circumvention activities unlikely to be captured by the scope of the intended studies by the Librarian.

218. *See supra* notes 24 and accompanying text. *See also* Benkler, *supra* note 24, at 416.

219. *See supra* notes 28-36 and accompanying text.

220. *See* 17 U.S.C.A. § 1201 (a)(1)(C) (West Supp. 1999) (setting forth factors); *see also* Benkler, *supra* note 24, at 420 ("[E]nforcement of the anti-device provision is unconstitutional unless and until the Librarian makes a determination that no non-infringing uses will be adversely affected by utilization of technological protection measures.").

221. *See supra* notes 136-140 and accompanying text for examples of other potential deleterious effects.

proof needed to maintain an action for anti-circumvention violations,²²² along with the generous remedies the Act provides,²²³ may prove to be a magnet for firms seeking to challenge acts of circumvention or devices that might, for example, concern trade secrecy or computer hacking matters.²²⁴ As long as there is a plausible claim that some material being protected by the technical protection system has a modicum of creative content that would entitle it to copyright protection,²²⁵ any act of circumvention or tool to aid the circumvention might be challenged under section 1201. Such uses of the statute could make copyright law into a general purpose misappropriation law regulating computer security matters. Moreover, as Part VI has shown, section 1201 is so ambiguous and broad that it may wreak considerable havoc in the information technology field, harming competition and innovation in this important sector. For these reasons, a broad regular review of the anti-circumvention rules should be undertaken.

VIII. CONCLUSION

The WIPO Copyright Treaty provides a “predictable, minimalist, consistent and simple legal environment” that should promote global commerce in electronic information products and services, in line with objectives and principles announced in the Clinton Administration’s *Framework for Global Electronic Commerce*.²²⁶ As the principal leader in the treaty-making effort that led to conclusion of this treaty, the Clinton Administration deserves credit for promoting this policy initiative that promises to substantially benefit the U.S. digital economy industries.

In most respects, the legislation implementing the WIPO Copyright Treaty in U.S. law also conforms to Framework principles.²²⁷ The anti-circumvention provisions of the DMCA, however, do not. They are unpredictable, overbroad, inconsistent, and complex. The many flaws in this

222. See *supra* notes 174-176, 193-195 and accompanying text.

223. See 17 U.S.C.A. § 1203(b) (West Supp. 1999) (civil remedy provision).

224. This potential was recognized in the Congressional debate over the anti-circumvention rules. See 144 CONG. REC., H7096 (daily ed. Aug. 4, 1998) (remarks of Rep. Goodblatte). Although Rep. Goodblatte indicated that computer hacking statutes should be used to deal with computer hacking problems, there is no reason why someone injured by a computer hacker could not seek relief under § 1201.

225. See 17 U.S.C. § 102 (1994) (copyright protection subsists in all original works of authorship that have been fixed in a tangible medium of expression).

226. See FRAMEWORK, *supra* note 1, at 3.

227. See *supra* notes 70-72 and accompanying text.

legislation are likely to be harmful to innovation and competition in the digital economy sector, and harmful to the public's broader interests in being able to make fair and other noninfringing uses of copyrighted works. If these regulations are not as maximalist as those initially proposed to Congress, this is mainly due to Congress' heeding of concerns expressed by some of the leading firms of digital economy interests, rather than to the Administration's leadership.

In the U.S. Congress, as well as in Geneva during the diplomatic conference leading up to adoption of the WIPO Copyright Treaty, proposed anti-circumvention regulations have been contentious. Among the concerns expressed in both venues were these: the potential effect of such rules on public access to information and on the ability to make noninfringing uses of copyrighted works, and their potential effect on competition and innovation in the market for hardware and software products whose uses might include the bypassing of some technical protection systems.²²⁸ The diplomatic conference had the good sense to adopt only a general norm on circumvention, leaving nations free to implement this norm in their own way.²²⁹ Thus, the flaws in the DMCA's anti-circumvention provisions do not derive from the treaty, but rather from the bad judgment of the Administration and the major copyright industry groups that urged adoption of overbroad rules in the DMCA.

This article has demonstrated that the DMCA's ban on the act of circumventing access controls should have included a general purpose "or other legitimate reasons" provision because the seven exceptions built into the statute, while they respond to the main concerns identified in the legislative debate, do not exhaust the legitimate reasons to bypass access controls.²³⁰ The article has provided examples of other legitimate circumvention activities and has suggested that if Congress does not narrow the reach of this provision, courts likely will do so, even if it involves some stretching to do so.

The article has also demonstrated that the anti-device provisions of the DMCA are substantially overbroad and need to be revised. The principal intent of Congress was to ban the development and deployment of "black boxes" that promote and enable piracy of copyrighted works.²³¹ However, the ban is far broader than this and threatens to bring about a flood of liti-

228. See *supra* notes 51, 87-89 and accompanying text.

229. See *supra* note 51 and accompanying text.

230. See *supra* notes 136-146 and accompanying text.

231. See *supra* notes 93 and accompanying text.

gation challenging a broad range of technologies, even where there is no proof that the technologies have or realistically would be widely used to enable piracy.²³² The legislation is also unclear about a crucial question: whether it is lawful for people to develop or distribute technologies that will enable implementation of the exceptions and limitations on the circumvention ban built into the statute.²³³ Did Congress intend to allow people to exercise these privileges, or did it intend to render these privileges meaningless because the technologies to enable the excepted activities have been made illegal? No clear answer to this question emerges from a careful study of the anti-circumvention provisions. While legislative clarification of this issue would be desirable, most likely the courts will have no choice but to address this question. Because of ambiguities in the statute, it is unclear how courts will resolve disputes in which such questions will be posed.

Finally, this article urges that the anti-circumvention provisions be subject to periodic interagency review that would consider their impact as a whole.²³⁴ The DMCA includes a provision authorizing the Librarian of Congress to study the impact of the act-of-circumvention provision and make a determination about whether this provision interferes with the ability of certain classes of users to make noninfringing uses of certain classes of copyrighted works.²³⁵ This provision is too narrow in at least two respects. One is that it does not perceive the potential impact of the device bans on the ability of users to make noninfringing uses of copyrighted works. The Librarian of Congress can and should consider this as well.²³⁶ A second is that the DMCA's study provision does not recognize other kinds of potential harm that the anti-circumvention provisions may do to competition and innovation in the information technology sector.²³⁷ Because of the unprecedented character of the anti-circumvention provisions and their overbreadth, it would be highly desirable for a broader study to be undertaken of the impact of these regulations with an eye to recommending changes to remedy unintended harmful consequences they may be having.

Before concluding this article, it is perhaps worth noting that as yet relatively few copyrighted works are being distributed with technical pro-

232. *See supra* notes 194-195 and accompanying text.

233. *See supra* notes 150-151 and accompanying text.

234. *See supra* notes 215-225 and accompanying text.

235. *See* 17 U.S.C.A. § 1201(a)(1)(B).

236. *See supra* notes 220 and accompanying text.

237. *See supra* notes 217 and accompanying text.

tection systems built in.²³⁸ Much research and development work is, however, underway to develop such systems.²³⁹ Many copyright owners seem to hope or expect that such systems will be widely used for a broad range of work in the not-too-distant future and that these systems will stop piracy and other unauthorized and arguably unlawful uses of copyrighted works.²⁴⁰

One factor that will significantly affect how widely technical protection systems will be deployed and how tightly they will restrict uses of copyrighted works is how consumers will react to them.²⁴¹ Copyright owners may feel far more secure if their works are technically protected so well that no unauthorized uses can ever be made of them. However, economists Carl Shapiro and Hal Varian argue that copyright owners must consider this:

The more liberal you make the terms under which customers can have access to your product, the more valuable it is to them. A product that can be shared with friends, loaned out and rented, repeatedly accessed, or sold in a resale market is obviously more valuable to a potential user than one that can be accessed only once, under controlled conditions, by only a single party.²⁴²

Moreover, people are very used to being able to make a wide range of uses of copyrighted works without seeking the copyright owner's permission. It is unclear how well they will react to a radical shift in the market for information products. Professor Benkler observes that "[w]e have no idea how a world in which information goods are perfectly excludable—as technical protection measures promise to make them—will look. Because

238. See COMPUTER SCIENCE AND TELECOMMS. BD., NATIONAL ACADEMY OF SCIENCES, *INTELLECTUAL PROPERTY RIGHTS AND THE EMERGING DIGITAL ECONOMY* (forthcoming 1999).

239. See Eric Schlachter, *The Intellectual Property Renaissance in Cyberspace: Why Copyright Law Could Be Unimportant on the Internet*, 12 BERKELEY TECH. L.J. 15, 38-45 (discussing various kinds of systems).

240. See Charles Clark, *The Publisher in the Digital World*, in *INTELLECTUAL PROPERTY RIGHTS AND NEW TECHNOLOGIES: PROCEEDINGS OF THE KNOWRIGHT'95 CONFERENCE 85* (Klaus Braunstein & Peter Paul Sint eds., 1995). See also White Paper, *supra* note 15, at 177-90 (foreseeing wide deployment).

241. Carl Shapiro and Hal Varian assert that "[t]rusted systems, cryptographic envelopes, and other copy protection schemes have their place but are unlikely to pay a significant role in mass-market information goods because of standardization problems and competitive pressures." CARL SHAPIRO & HAL VARIAN, *INFORMATION RULES* 102 (1998).

242. *Id.* at 98.

of the non-rival nature of information, prevailing economic theory would suggest that we are as likely to lose as gain productivity from this technological change."²⁴³ In addition, if consumers won't buy tightly restricted copies, copyright owners may end up worse off than before.²⁴⁴

Competition among information providers may also affect the successful deployment of technical protection systems. If one information provider tightly locks up his content, a competing provider may see a business opportunity in supplying a less tightly restricted copy to customers who might otherwise buy from the first provider.²⁴⁵ A competitive alternative to tight technical controls may well be to adopt one of the several strategies that Shapiro and Varian discuss to show how content providers can take advantage of the opportunities presented by digital technologies, rather than being overwhelmed by the risks.²⁴⁶ There are, they say, many other good business models out there waiting to be invented by creative content providers.²⁴⁷

If content providers come to believe that a good business model is the best way to protect intellectual property from market-destructive appropriations, perhaps the current debate over the DMCA's anti-circumvention regulations will seem in time like a tempest in a teapot. However, in the meantime, the impact of this legislation should be closely watched because of its potential for substantial unintended detrimental consequences.

243. Benkler, *supra* note 24, at 424.

244. See Branko Geravac et al., *Electronic Commerce and Intellectual Property—Protect Revenues, Not Bits*, 2 IMA INTELL. PROP. PROC. 111 (1996).

245. This, in essence, is what happened when software developers, such as Lotus Development Corp. started distributing copy-protected versions of their programs. Firms with similar products who were willing to sell their products without copy-protection systems attracted enough customers that the leading firms eventually abandoned their technical protection schemes. This is, of course, more likely to occur where markets are competitive and where participants in the market are not acting jointly in deciding on technologies so that consumers will not have a competitive choice.

246. See SHAPIRO & VARIAN, *supra* note 241, ch. 4.

247. See *id.* at 84. Some of these business models may themselves be subject to intellectual property protection. See, e.g., Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts*, 14 BERKELEY TECH. L.J. 577 (1999).

COMMENTARY: BLACK HOLES OF INNOVATION IN THE SOFTWARE ARTS

By Mark A. Haynes[†]

I. INTRODUCTION

In order to promote technological progress, the intellectual property system relies on a careful balance between free access to information about products on the market on the one hand, and the encouragement of investment in innovation through the exclusive rights of patent, copyright, and trade secret laws on the other hand. In the software arts, including e-commerce, the system is unbalanced by the tendency of copyright law to isolate technology in pockets controlled by the original author, and to thus act as the gravity for "black holes" of innovation.

The Magaziner Report recommended that strong patents in the area of e-commerce would further encourage technological development,¹ and some steps have been taken to that affect. The courts have recognized the patentability of software-based inventions,² and the software elite have begun to participate in the patent system.³ As a result, the volume of patent applications being filed in the areas relating to e-commerce is exploding.

However, simply increasing the number of patents in software technologies does not ensure that the goal of promoting progress in the technology is achieved. Copyright laws need to be reevaluated and modified to

© 1999 Mark A. Haynes.

[†] B.S.E.E., 1977, University of Texas at Austin.; J.D., 1981, Stanford Law School. Admitted to practice before the U.S. Patent and Trademark Office, 1982. The author is the founding partner of Haynes & Beffel LLP, a patent prosecution firm in Half Moon Bay, California. The author would like to thank the editorial staff of the *Berkeley Technology Law Journal* for their assistance.

1. See WILLIAM J. CLINTON & ALBERT GORE, JR., A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE § 3 (1997), available at <<http://www.iitf.nist.gov/eleccomm/ecommm.htm>>.

2. See, e.g., *Diamond v. Dierh*, 450 U.S. 175 (1981); *In re Meyer*, 688 F.2d 789 (C.C.P.A. 1982); *Arrhythmia Research Tech., Inc. v. Corazonix Corp.*, 958 F.2d 1053 (1992).

3. See, e.g., Rodney Ho, *Patents Hit Record in '98 as Tech Firms Rush to Protect Intellectual Property*, WALL ST. J., Jan. 15, 1999, at A2; Jonathan M. Moses, *When Copyright Law Disappoints, Software Firms Find Alternatives*, WALL ST. J., May 4, 1993, at B6.

create an appropriate balance in intellectual property laws to promote technological progress in the software arts. In particular, I believe that rules enabling reverse engineering of software-based products should be revisited and amended to allow competing companies and individuals to study each other's work in detail and use the ideas discovered without fear of copyright liability. The reverse engineering right in the Semiconductor Chip Protection Act of 1984⁴ may provide a suitable model for legislation in this area.

This paper comments on the role copyright plays, independently and in concert with patents, in blocking innovation in the software arts. Part II describes how copyright slows the pace of innovation in the software arts, particularly by removing much unpatented but important technology from the public domain and preventing effective reverse engineering. Part III discusses how the patent system fuels innovation, and argues that copyright upsets the balance in intellectual property. The paper concludes with a call for further study on the effect of copyright on innovation in the software industry, and legislative reform to further enable reverse engineering.

II. HOW COPYRIGHT SLOWS THE PACE OF INNOVATION IN THE SOFTWARE ARTS

Many comparisons have been made between the programming art and the arts traditionally covered by copyright, such as writing and painting.⁵ These comparisons feed the argument that intellectual property law should focus on the protection of developers of code from others who might pick-

4. The Semiconductor Chip Protection Act of 1984, Pub. L. No. 98-260, 98 Stat. 3335, 3347 (codified as amended at 17 U.S.C. §§ 901-14). The reverse engineering provision of the Act provides that it is not an infringement of the exclusive rights of the owner of a mask work for:

(1) a person to reproduce the mask work solely for the purpose of teaching, analyzing, or evaluating the concepts or techniques embodied in the mask work or the circuitry, logic flow, or organization of components used in the mask work; or

(2) a person who performs the analysis or evaluation described in paragraph (1) to incorporate the results of such conduct in an original mask work which is made to be distributed.

17 U.S.C. § 906(a) (1994).

5. See, e.g., Anthony L. Clapes, *Confessions of an Amicus Curiae: Technophobia, Law and the Creativity in the Digital Arts*, 19 DAYTON L. REV. 903 (1994); Arthur R. Miller, *Copyright Protection for Computer Programs, Databases, and Computer-Generated Works: Is Anything New Since CONTU?*, 106 HARV. L. REV. 977 (1993).

up and improve their work, as if software developers were artists working in isolation.⁶ But it is just this isolation of software technology, induced by copyright, that damages the processes of innovation. The isolation of the developer is important for the content, as opposed to the technology, of e-commerce. Unlike the content on the web, the technology for delivering content needs to be shared and understood by many in order to advance. Copyright laws must be understood to determine how they fuel, or block, the free flow of information.

My experience as a patent attorney, coupled with anecdotal examples from legal scholarship and the market, leads me to believe that software engineers are constantly reinventing the wheel (and patenting old ideas) because key avenues for learning what others have already done are not open to them. Notwithstanding the creativity of individuals in the field, as long as they are blocked from careful study of the work of others much of the creativity is wasted traversing ground exhaustively explored by others before them.⁷ Innovation is slowed because creative energies are applied to ignore the work of others and to re-do basic designs in an effort to avoid accusations of copyright infringement, rather than to build on and improve the state of the art, which necessarily includes the work of others.

This Part discusses how copyright effectively blocks important “conduits” that facilitate the flow of information into the marketplace. We can see the effects of this restriction in several (admittedly anecdotal) real-world examples. In particular, the open source movement can be understood in this context as a way to innovate around the stifling effects of copyright on innovation.

A. Copyright Restricts the Flow of Information

Innovation is fed by three key conduits for the flow of information: (1) reverse engineering of products, (2) patents, and (3) technical literature. However, copyright blocks each of these conduits of information. In doing so, copyright essentially provides the gravity for black holes in which no one but the copyright holder is able to innovate, effectively blocking competition for compatible products, and isolating the technology from improvement by anyone but the copyright holder.

First, copyright essentially takes away the freedom to reverse engineer. In fact, with only very limited exceptions, it is copyright infringement to

6. *See id.*

7. *See* Dennis S. Karjala, *Copyright Protection of Computer Documents, Reverse Engineering, and Professor Miller*, 19 DAYTON L. REV. 975 (1994).

copy someone else's code to study it, and to apply the knowledge gained to make competing products, under the so-called "intermediate copying" rules.⁸ One defense to copyright infringement is "fair use" under Section 107 of the Copyright Act.⁹ However, as discussed in more detail below, it offers a poor refuge to those seeking to reverse engineer software to derive the functional aspects. Because fair use is an affirmative defense, the burden is on the person reverse engineering the product to prove fair use, and this alone serves as a deterrent. The defense is further weakened by the presumption that commercial uses are unfair.¹⁰

Additionally, while two Court of Appeals decisions recognized a limited right of fair use, they were decided on very narrow grounds, and therefore are not much help. In *Atari Games Corp., v. Nintendo of America*, the Federal Circuit recognized that copying and reverse engineering are not copyright infringement when the copying and reverse engineering are necessary to learn the unprotected ideas and processes.¹¹ The Ninth Circuit decision in *Sega v. Accolade* similarly found that copying and reverse engineering computer programs can be a fair use under the Copyright Act.¹² However, the holding in *Sega* only applies in the narrow circumstance of when a programmer needs to reverse engineer in order to ensure compatibility with his own product, and no other means of access exists.¹³ Although both of these decisions potentially allow a fair use defense for reverse engineering, practical realities make this extremely difficult; the ad hoc nature of copyright enforcement, the variability of outcomes, and the need to rely on an affirmative defense makes "fair use" a thin basis for competitors interested in learning the unprotected ideas and processes from computer programs through copying and reverse engineering. As a result, lawyers tell programmers to avoid looking at the work of others for fear of copyright infringement.

8. See *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992); *Atari Games Corp. v. Nintendo of Am.*, 975 F.2d 832 (Fed. Cir. 1992). See also John G. Mills, *Possible Defenses to Complaints for Copyright Infringement and Reverse Engineering of Computer Software: Implications for Antitrust and I.P. Law*, 80 J. PAT. & TRADEMARK OFF. SOC'Y, 101 (1998).

9. 17 U.S.C. § 107 (1994).

10. See *Sega Enters.*, 977 F.2d at 1522.

11. See *Atari*, 975 F.2d at 844. Note that the Federal Circuit found that Atari had exceeded the scope of that right. See *id.*

12. See *Sega Enters.*, 977 F.2d at 1520-21.

13. See *id.* at 1521. Unlike in the Atari case, the Sega court found Accolade's intermediate copying to be a fair use. See *id.*

Second, copyright reduces the incentive for software companies to patent their products. The exclusivity provided by patents is expensive, requires a degree of value added to the state of the art, and can be neutralized by improvement patents. In contrast, exclusive copyrights come at virtually no cost to the holder. With fewer patents, companies keep more information about important ideas secret and avoid gauging their work against that of others. So far, patents describe only a small part of the software arts. But many software companies are now quite aggressively applying for patents.¹⁴

Finally, without free reverse engineering, we are left with the technical literature as a means for sharing ideas in the software arts. But technical publications are generally theoretical in nature, and very shallow compared to real products. Without reverse engineering, there are huge gaps in usable information.

B. Evidence of the Slow Pace of Innovation

To verify the harm copyright has done to the software field by restricting these conduits of information, what we need is a comparative review of the processes of innovation in the software field with other high technology fields not protected by copyright. The goal of this review would be two-fold: (1) to provide insights needed to sort out the intellectual property laws related to computer software, and (2) to restore a balance that promotes innovation and fair competition in the software field, like the balance achieved in other fields of high technology.

In the absence of such a study, anecdotal evidence must suffice. Commentary from the legal community on intellectual property laws as applied to computer software includes some interesting perspectives on the lack of invention in the software field. For example, it has been argued that "innovation in the software field is primarily incremental. Invention in the formal sense of the term is rare."¹⁵ To a patent attorney, this is an oxymoron. Innovation *is* invention. Practically all inventions in all fields are incremental. Inventions are *patentable* only if they have not been made before by another.

Yet another commentator has argued that constant reinvention in the software field suggests that patents are not needed.¹⁶ After all, if these in-

14. See Ho, *supra* note 3; Moses, *supra* note 3.

15. Pamela Samuelson et al., *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308, 2376 (1994).

16. See John Swinson, *Copyright or Patent or Both: An Algorithmic Approach to Computer Software Protection*, 5 HARV. J. L. & TECH. 145, 157 (1991).

novations can be so readily remade, there is no reason to encourage dissemination of them, or so the argument goes. This, of course, is nonsense. Re-making of inventions wastes resources, and slows down the art. In the software field, the programmers are constantly going over old ground, carefully avoiding access to the code of their competitors, and re-inventing. Access to the work of others prevents this wasteful effort.

Comments like these may actually be a reflection of the very slow pace of patentable innovation caused by copyright in the software field. If an inventor must start from scratch every time a new product is developed, unable to study and learn from the work of others, it is very likely the inventor will not be using the optimal approach, since he cannot learn from others what is optimal. This taps intellectual resources, and slows down development of the art.

The most obvious evidence of slow moving technology attributable to copyright is Microsoft's Windows operating system. The copyright on Microsoft Windows blocks the flow of information needed for innovation in the field of compatible personal computer operating systems by would-be competitors. In addition, as the operating system grows through creation of derivative works, the copyright expands. Microsoft is able to exploit the imbalance in the intellectual property laws to expand its copyrighted franchise, and to build a monopoly that naturally expands with the copyright.¹⁷

The brake on innovation that copyright engenders can be seen by simply comparing the rate of new product releases for Windows to the rate for x86 microprocessors. When Windows 95 came out, it was considered a technological non-event. When Windows 98 was released, the primary innovation in this version that people were talking about was the integration of the Internet browser with the operating system. Commercial x86 microprocessors made by Intel and others, on the other hand, have gone through many revisions since 1995. Intel has enough competition in compatible microprocessors from AMD, National Semiconductor, and others to drive innovation.¹⁸ This competition is able to flourish in part because there are no fundamental copyrights blocking access to the microprocessor

17. See Mark A. Lemley, *Antitrust and the Internet Standardization Problem*, 28 CONN. L. REV. 1031, 1053-54 (1996). See also Robert H. Lande & Strugis M. Sobin, *Reverse Engineering of Computer Software and Antitrust Law*, 9 HARV. J. L. & TECH. 237 (1996).

18. See Luc Hatlestad, *Finally Some Competition, AMD and Others are Making a Legitimate Run at Intel*, THE RED HERRING, Apr. 1998, available at <<http://www.redherring.com/mag/issue53/competition.html>>.

art. Generally, innovators in the software arts are working on products other than the most widely used operating system. This slows down the advancement of the technology to a snail's pace.

Other evidence of the slow pace of innovation in the software field is found in the attacks made on the patents that are issuing for software-based inventions. The software industry loudly complains that software patents are being granted for ideas that are old. But why would a software developer believe that his or her idea was patentable and invest thousands of dollars to get a patent if he or she thought it was an old idea? The only answer to this question that is not cynical is that the work of others is not available to the inventors for study.

The "open source" movement in the software field, viewed in this light, is a clear reaction to the closed nature of copyright protected software.¹⁹ According to this movement, software should be distributed with a liberal copyright waiver, and source code made available to anyone that desires to modify or improve it. With open source, products evolve with input from many developers who all benefit from the work of earlier contributors. Evolution leads to efficient, stable, and versatile products. This movement has gained momentum lately with such acclaimed products as the LINUX operating system, the Apache web server software, the PERL scripting language, and a popular e-mail server product called Sendmail.

However, even the open source movement exploits and is hindered by copyright. The most popular open source code products (LINUX and others) are distributed under the General Public License of the Free Software Foundation. The copyright license limits use of the open source code to products that are also distributed as open source, and for which no patent licenses are required. This aspect of open source distribution agreements has been described as "viral."²⁰ Thus, users of open source risk losing their own copyrights and contract the patent system out of the product. They lose the benefits of the intellectual property system designed to protect their investment.

III. THE PATENT SYSTEM CONTRASTED

To my understanding, the software industry is the only one that requires such a movement to enable an understanding of the work of others

19. See Tim O'Reilly, *The Open Source Revolution*, RELEASE 1.0, Nov. 1998.

20. *Id.*

to promote the advancement of technology. The patent system, by contrast, fuels innovation in fields that lack exposure to copyright.

Unlike copyright, the patent system encourages improvement patents, through which competitors are able to neutralize the patent portfolios of others. People who contribute to the advancement of the technology build relevant patent portfolios to trade with one another. Innovators therefore have freedom of action to utilize the best available technology for their products, whether patented by them or by others. In this way, patents ensure that competitors do the work to advance the technology because only contributors obtain relevant, tradable patents. Those that take a free ride do not obtain patents to trade and become exposed to the patents of others to a much greater degree. Free riders usually end up paying greater license fees or being pushed out of the market.

Because patents are based on claims that define the protected technology reasonably well, competitors can often avoid infringement. The process of avoiding patents, more often than not, results in improvements and new thinking, and thereby accelerates the growth of the technology. Compare Intel in the patent-dominated microprocessor art and Microsoft's operating system copyright. Intel introduces improved x86 processors at frequent intervals. Microsoft made minor upgrades to Windows in 1995 and 1998, but Microsoft does not seem to be in any hurry to introduce any new operating systems. Intel is driven to innovate by the patent system, reverse engineering, and the lack of core copyrights that enable innovative competitors like AMD and National Semiconductor to make and sell compatible products. The slower Intel innovates, the more market share it loses to the companies nipping at its heels. In contrast, the copyright on Windows has stopped all would-be competitors cold. In the trenches, no one is working hard to improve the core of Windows—at least not the way Intel, AMD, and National Semiconductor work hard to improve the microprocessor. Microsoft, however, has not suffered the stress of competition needed for real innovation at the core of the operating system. This can be explained in large part by the imbalance in incentives created by copyright.

Copyright tends to protect the accumulation of simple design choices made by programmers, leaving only the fundamental ideas represented by their work in the public domain. Patents, on the other hand, tend to protect ideas that inventors think are fundamental, leaving everything else in the public domain. This contrasting approach to protection has led some to fear patents in the software field. But the evidence in other fields does not bear out this fear of patents. One learns in practice that the patent system is not so much about protecting what is patented as it is about leaving the

unpatented in the public domain. Copyright removes much unpatented but important technology from the public domain, and upsets the balance of the intellectual property system.

IV. CONCLUSION

Microsoft's monopoly, and other smaller, technical black holes fueled by copyright, like the Apple Macintosh operating system or the WordPerfect word processor, are bad for innovation. We should limit the scope of the copyright in software to enable reverse engineering not only for access to the information in the software, but for use of the information to make compatible products.

Perhaps the approach taken in the Semiconductor Chip Protection Act of 1984 could work in the case of software arts. In that Act, normal infringement was to be found based on substantial "similarity," like the copyright test.²¹ However, if the accused could prove that legitimate reverse engineering was conducted, the infringement would be found using a stricter substantial "identity" test.²² If we made such changes, the industry might have the information necessary for advancement, and the freedom to use it. Someone would find a way to push Microsoft to innovate in its core technology, like AMD and National Semiconductor push Intel. If consumers could buy a Windows 95 compatible alternative from someone else, Microsoft's Windows 98 would surely be a real improvement. Microsoft would be forced to invent in order to maintain market share. The software industry would benefit, and the constitutional mandate to "promote the Progress of ... useful Arts" would be served.²³

21. See *Brooktree Corp. v. Advanced Micro Devices, Inc.*, 977 F.2d 1555, 1564 (9th Cir. 1992).

22. See *Brooktree Corp. v. Advanced Micro Devices, Inc.*, 705 F. Supp. 491, 495 (S.D. Ca. 1988).

23. U.S. CONST. art. I, § 8, cl. 8.

AS MANY AS SIX IMPOSSIBLE PATENTS BEFORE BREAKFAST: PROPERTY RIGHTS FOR BUSINESS CONCEPTS AND PATENT SYSTEM REFORM

By Robert P. Merges[†]

ABSTRACT

In this paper, Professor Merges describes the emergence of patents for business “methods” or concepts, such as Internet airplane ticket purchase systems. Professor Merges is agnostic about whether these patents are worthwhile. Nevertheless, he argues that the increased volume of patent applications stemming from this newly patentable subject matter has pushed the patent system into crisis. In particular, he focuses attention on determining an acceptable “error rate” for issued patents, with an eye toward reducing the number of invalid business concept patents that are actually issued. In the process, he calls for new appreciation of the relationship between the patent office and private parties. He argues for policies that will efficiently coordinate the efforts of both groups to achieve the socially desirable end, which is an appropriate expenditure to determine patent validity. Some of these reforms involve restructuring jobs and incentives in the Patent Office. Others involve obtaining the input of those parties that suffer most if a firm receives an invalid patent—i.e., the firm’s competitors. These also tend to be the parties with the best information about patent validity. It is therefore logical, according to Professor Merges, to get these competitors into the patent process as early and as thoroughly as possible. This leads to a proposal to adopt a patent opposition system in the U.S., much like the one currently in place in Europe. Only reforms such as these will lower the incidence of poor-quality patents. And only then, Merges argues, will we be able to decide whether patents for business concepts make sense or not.

TABLE OF CONTENTS

| | |
|--|-----|
| I. INTRODUCTION | 578 |
| II. BACKGROUND: THE “IMPOSSIBLE” IS NOW POSSIBLE | 581 |
| III. HOW WE GOT TO WONDERLAND | 584 |
| IV. EVALUATING THE PATENT EXAMINATION SYSTEM..... | 588 |

© 1999 Robert P. Merges.

[†] Wilson Sonsini Goodrich & Rosati Professor of Law, Boalt Hall School of Law, University of California, Berkeley.

| | | |
|-----|--|-----|
| A. | Why Is Patent Quality So Poor? | 589 |
| B. | Sketching an Ideal Patent Office..... | 591 |
| | 1. <i>The function of a patent office</i> | 591 |
| | 2. <i>The goals of a patent office</i> | 592 |
| | 3. <i>Pros and cons of a simple registration system</i> | 594 |
| C. | Optimal Public Expenditures on Patents..... | 596 |
| | 1. <i>The benefits of sorting applications</i> | 596 |
| | 2. <i>The problems (and politics) of sorting</i> | 597 |
| | 3. <i>Second-best solutions</i> | 598 |
| | a) Why not penalize holders of invalid patents? | 599 |
| | b) Who is the cheapest cost-avoider? | 599 |
| D. | Examining the PTO | 600 |
| | 1. <i>The PTO's examination budget</i> | 600 |
| | 2. <i>Setting the ideal PTO budget</i> | 603 |
| V. | SOME SIMPLE SUGGESTIONS FOR IMPROVING THE EXAMINATION PROCESS | 606 |
| | A. Job Design | 606 |
| | B. Alternative Bonus Systems | 609 |
| | C. Reforming Reexaminations: The Common Sense Case for Patent Oppositions | 610 |
| VI. | CONCLUSION..... | 615 |

"Now I'll give you something to believe!" the White Queen remarked.] "I'm just one hundred and one, five months and a day."

"I can't believe that!" said Alice.

"Can't you?" the Queen said in a pitying tone. "Try again, draw a long breath and shut your eyes."

Alice laughed. "There's no use trying," she said, "one can't believe impossible things."

"I daresay you haven't had much practice," said the Queen. "When I was your age, I always did it for half-an-hour a day. Why sometimes I've believed as many as six impossible things before breakfast."¹

I. INTRODUCTION

The White Queen would be right at home in the U.S. patent system today. First software, once thought too purely mathematical, and now business "methods" or concepts, once thought too abstract, have become per-

1. LEWIS CARROLL, *THE ANNOTATED ALICE* 250-51 (Martin Gardner, ed. 1960). This passage comes by way of political scientist Don Herzog. See Don Herzog, *As Many as Six Impossible Things Before Breakfast*, 75 CALIF. L. REV. 609 (1987) (critiquing "Critical Legal Studies"), quoting from LEWIS CARROLL, *THROUGH THE LOOKING GLASS* (1871).

fectly acceptable subject matter for patents. For better or for worse, whole new landscapes have been opened to the possibility of patents.

To get right to the heart of the issues surrounding patents for business concepts, log on to <http://www.walkerdigital.com/html/information.html>. This is the website of Walker Digital, Inc., the company that recently “spun off” its Priceline.com subsidiary, a separate company that uses the Internet to match buyers with sellers.² Here is what you will read:

We [Walker Digital, Inc.] conceive, research, and prepare our patented business systems in-house. Our team of specialists prepares cases that solve real-life problems for a wide variety of industries such as retail, telecommunications, credit cards, casinos and more. So far, we’ve filed over 250 U.S. and international patent applications to create a portfolio that we believe is unlike anything anywhere else in the world.³

Until very recently, Walker Digital would not have existed. The patent system did not embrace the abstract patents on business concepts that are the company’s key assets.⁴ There would be no cornerstone patents on internet price-matching, personified by Walker Digital’s “Priceline.com” subsidiary.⁵ Without patents, in fact, it is difficult to see how a firm could survive as an independent “idea factory” for Internet commerce.

2. See generally Priceline.com, *Priceline.com* (visited Apr. 19, 1999) <http://www.priceline.com/>.

3. Walker Digital Corp., *Information* (visited Apr. 24, 1999) <http://www.walkerdigital.com/html/information.html>.

4. See *State Street Bank & Trust Co., Inc. v. Signature Financial Group, Inc.*, 149 F.3d 1368 (Fed Cir. 1998) (overruling cases holding or suggesting that claims to “methods of doing business” were not patentable). In many ways, *State Street Bank* did not initiate a new practice; it lent judicial authority to existing PTO policy:

With regard to “methods of doing business” in particular, it is worth mentioning that there are a large number of patents in this category that have been granted by the U.S. Patent and Trademark Office (PTO) prior to the Federal Circuit’s *State Street* decision. In effect, *State Street* will serve to help confirm their validity.

Scott M. Alter, *Federal Circuit Broadens Scope For Software Patents*, 15 COMPUTER LAW. 24, 27 (1998).

5. Consider one line of business that Priceline.com is apparently interested in: airline ticket options, i.e., the purchase and sale of the right to buy tickets at a later time for a specified price. See U.S. Patent No. 5,797,127, issued Aug. 18, 1998 (entitled “Method, Apparatus, And Program For Pricing, Selling, And Exercising Options To Purchase Airline Tickets”). This patent has the two attributes of a business concept patent: (1) it describes an essentially *commercial* (as opposed to technological) activity, typically some

Walker Digital is therefore a perfect test case. It can tell us whether formerly "impossible" patents on business concepts are a good idea. If there were some way to determine whether this firm had initiated business concepts that no one else would have, or had hurried them into practice faster, we could ask: is the game worth the candle? Alas, no such knowledge has been revealed to us. The instruments we have at hand are simply too imprecise, at least for the time being. We may see an explosion of activity. Or we may hear horror stories about good, solid businesses aban-

way to make or save money; and (2) the hardware and software elements are described and claimed at such a high level of generality that they are for all practical purposes nominal. These features are readily apparent from the abstract and claim 1:

An apparatus, method, and program for determining a price of an option to purchase an airline ticket, and for facilitating the sale and exercise of those options. By purchasing an option, a customer can lock in a specified airfare without tying up his money and without risking the loss of the ticket price if his travel plans change. Pricing of the options may be based on departure location criteria, destination location criteria, and travel criteria.

[Claim 1:]

A data processing apparatus for determining a price of an option to purchase an airline ticket, comprising:

 a central controller including a CPU and a memory operatively connected to said CPU;

 at least one terminal, adapted for communicating with said central controller, for transmitting to said central controller option pricing information including departure location criteria, destination location criteria, and travel criteria;

 said memory in said central controller containing a program, adapted to be executed by said CPU, for calculating a price of an option to purchase within a future period, for a particular ticket price, an airline ticket satisfying the departure location criteria, destination location criteria, and travel criteria;

 wherein said central controller receives said criteria from said terminal and calculates the option price based upon the criteria.

The emphasis on the *commercial function* of the program ("calculating a price of an option to purchase ... an airline ticket"), together with the complete generality of the hardware and software elements ("central controller," "at least one terminal," "CPU," "memory," and "a program" are all completely general), leads to the conclusion that this is a patent on the business idea of using computers, in particular the Internet, to price and purchase options on airline tickets. For other examples of patents such as this, see U.S. Patent No. 5,732,400, issued Mar. 24, 1998 (entitled "System And Method For a Risk-Based Purchase Of Goods"); U.S. Patent No. 5,787,402, issued July 28, 1998 (entitled "System and Method for Performing Automated Financial Transactions Involving Foreign Currencies").

done in the face of predatory patent extortionists. It is simply too soon to tell.

But there *are* some positive steps we can take to limit any negative effects from business method patents. The most important is to make sure that the business concept patents that do issue are good, solid patents. It may be too late to argue to a court that business concept patents are universally bad. And it may be too early to ask Congress to rein them in. But it is neither too late nor too early to argue forcefully that *bad* business concept patents are bad.⁶ In fact, the time is just right: minimizing the number of worthless business concept patents makes a great deal of sense just now. Only by improving the overall quality of these patents can we begin to determine whether or not they make any sense. Once we disentangle the bad from the good, we can see whether the good ones are worth the trouble. If, in the process, this entails improving the overall quality of issued patents, all the better. If it tweaks us into fixing some deep-seated flaws in the way the Patent and Trademark Office ("PTO") examines patents, the advent of business method patents may even turn out to serve a useful purpose.

II. BACKGROUND: THE "IMPOSSIBLE" IS NOW POSSIBLE

Before addressing the question of bad business concept patents, let us first consider how we came to patent this subject matter in the first place. Although the older cases do not articulate their reasoning very clearly, they seem to center around one idea: that the patent system was meant to protect *technology*—actual machines, devices, and new chemical compositions—rather than pure concepts.⁷ Because business methods are not tied to particular machinery or devices, they are clearly not patentable under this view.

This antipathy to patenting mere abstractions actually grew out of older cases which questioned the patentability of processes per se.⁸ How,

6. It will become clear as I go along what I mean by a "bad" patent. Succinctly put, it means a patent that should have been weeded out after a reasonable investment of effort, but was not.

7. See, e.g., *Gottschalk v. Benson*, 409 U.S. 63 (1972). See generally ROBERT P. MERGES, *PATENT LAW AND POLICY: CASES AND MATERIALS* ch. 2 (2d ed. 1997).

8. This history, which culminated in the acceptance of process patents in *Cochrane v. Deener*, 94 U.S. 780, 788 (1877), is well recounted in DONALD CHISUM, *CHISUM ON PATENTS* § 1.03 (1978 & Supp. 1999). See generally Edward C. Walterscheid, *The Early Evolution of the United States Patent Law: Antecedents (Part 3 Continued)*, 77 J. PAT. & TRADEMARK OFF. SOC'Y 847 (1995).

it was asked, could a list of steps not tied to particular machinery or devices be patentable? In time, the opposition to process patents died away, partly because they came to be understood as *physical transformations* rather than mere abstractions.⁹ It also did not hurt that they were perceived as crucial to the growing chemical industry of the early twentieth century. Yet, the prohibition on patents for business methods lived on.¹⁰

With the acceptance of patents for software, courts could no longer persuasively rely on the distinction between concepts and machines.¹¹ Even so, for a brief time the rule against business method patents survived. Those who defended this rule justified it on other grounds. Most powerfully, it was argued that such patents were simply not necessary.¹² After all, there seemed to be no shortage of new accounting methods, financial instruments, or financial services techniques throughout the history of the American economy, when business methods were not patentable. Even into the mid-1980s, when business method patents were just beginning to appear, the U.S. was considered the world leader in this service industry. Thus, according to this view, the proper question is: why fix it if it ain't broke?

The conventional answer is dictated by the logic of patent principles and current practices. It holds that there is no sound reason *not* to protect business methods. The history, logic, and accepted practices of our method of granting patents essentially compels us to allow patents on business concepts, because there is no principled basis on which to distinguish this "industry" from the myriad other industries that routinely obtain patents. Further, we should all have faith that this wave of patenting will unleash an Edisonian tidal wave of inventiveness—that, if we thought entrepreneurs rapidly introduced new ideas such as overnight package deliv-

9. See *Cochran*, 94 U.S. at 780.

10. See, e.g., *Hotel Security Checking Co. v. Lorraine Co.*, 160 F. 467, 469 (2d Cir. 1908).

11. They certainly tried, nonetheless: See, e.g., *In re Alappat*, 33 F.3d 1526 (Fed. Cir. 1994) (en banc) (emphasizing hardware components of claimed "rasterizer" invention). But see Pamela Samuelson, *Benson Revisited: The Case Against Patent Protection for Algorithms and Other Computer Program-Related Inventions*, 39 EMORY L.J. 1025 (1990) (making strong case against software patents).

12. See MERGES, *supra* note 7, at 156 ("Regardless of specific strategies, the point is the same: firms can capture the value of innovations many ways. The question for policymakers is whether patents should be permitted, in light of the other "appropriability mechanisms" available. Again, the relatively frequent innovations in the financial services industry prior to the era of patentability suggest that firms had adequate means to appropriate the value of their new financial innovations.").

ery and 1-800-Flowers *without patents*, then Watch Out!, because we haven't seen *anything* yet in this field!

Certainly Walker Digital sees it this way. Again, their web page:

We're not believers in traditional commercial inventing, where old methods are shoehorned into new technologies. Rather than think outside the box, we seek to reinvent the box. We create practical new ways to do things based on the inherent benefits of new technologies. We then take our core ideas, protect them with patents and establish licensing partnerships with major industry players who bring our ideas to market. These ideas are our intellectual property, our product.

Who are we?

Our team includes entrepreneurs, inventors, technologists, patent attorneys, industry analysts and even a world-renowned cryptographer. It also includes folks who, in previous lives, were some of the country's top CEOs and marketing executives. It's a group of highly intelligent, inventive, business-savvy people, with plenty of room for even more bright people, like you.

We earn profits from our intellectual property through a variety of business strategies ranging from direct licensing agreements—selling an idea to another company—to spinning off new businesses in which we retain an equity stake. Our first business spin-off was a home run—Priceline.com, our patented buyer-driven commerce system.

Priceline.com is the only system, on or off the Internet, where buyers can name their own price for specified goods and services. Like all of our spin-offs, its success has become our success.¹³

Not surprisingly, the patent covering the Priceline.com service is in dispute. A rival inventor—a patent lawyer, in fact—had filed a patent application with somewhat analogous claims earlier. An interference is now afoot.¹⁴ And so commences the inevitable shakeout period when rival patentees jockey for position. This much, at least, is not new. For example, the Bell System had 600 patent infringement suits pending in the late

13. Walker Digital, *Information* (visited Apr. 24, 1999) <<http://www.walkerdigital.com/html/information.html>>.

14. See Teresa Riordan, *It May Be "Big, Really Big," But An On-Line Airline Ticket Discounter is Also Being Challenged*, N.Y. TIMES, Jan. 18, 1999, at C1. A patent interference is a proceeding to determine priority among two or more rival inventors. See 35 U.S.C. § 135 (1998).

nineteenth century,¹⁵ and the past fifteen years have seen a steady stream of foundational litigation in the biotechnology industry.¹⁶

What *is* new is this: the shakeout has begun without answers to some important threshold questions. Chief among these is whether Walker Digital and other firms like it are doing anything that would not be done in the absence of patents. Put another way, in an ideal world, society would have addressed whether or not the types of business concept patents sought by these firms contributed any value in excess of what they cost society. If the answer was no, we would deny patents to them; if yes, patents would be allowed.

III. HOW WE GOT TO WONDERLAND

It would certainly be nice to have a theory that would tell us when one type of invention is unpatentable, while another is patentable. But the problem with such a normative theory of patentable subject matter is no less vexing for its familiarity. Where do you draw the baseline? One conventional candidate, historical practice, is not helpful. There would seem to be little hope in constructing an “originalist” interpretation of the Intellectual Property Clause of the Constitution¹⁷ to limit its subject matter: By definition, the clause envisions the creation of unanticipated inventions and writings. It provides no built-in limits. Hence it does little good to argue that the patent law traditionally protects only conventional “hardware” inventions. It is quite true that the canonical patented technology in the eighteenth century was a simple agricultural tool (an axe or a plow) which then became a more complex implement (a cotton gin or reaper) in the nineteenth century; even later, it became a machine, electrical device, or chemical process. These are true, but useless, historical facts; they say nothing about the appropriateness of patenting modern business concepts.

Indeed, following in the general spirit of the Intellectual Property Clause, Congress early on seems to have embraced a kind of blind technological optimism, ignoring economic costs. The value of a limited and well-administered patent system was little debated in the early Republic.

15. See DAVID NOBLE, *AMERICA BY DESIGN: SCIENCE, TECHNOLOGY AND THE RISE OF CORPORATE CAPITALISM* 10 (1977); JOHN BROOKS, *TELEPHONE: THE FIRST HUNDRED YEARS* 77 (1975).

16. See generally KENNETH BURCHFIELD, *BIOTECHNOLOGY AND THE FEDERAL CIRCUIT* (1995 & Supp. 1997).

17. U.S. CONST., art. I, § 8, cl. 8 (“The Congress shall have Power ... To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries...”).

There were no detailed cost-benefit analyses when the first few patent acts were passed. And there was consequently no real effort to separate patentable subject matter from things that were not to receive patents. Perhaps this was in part a function of an understanding—shared widely among legislators, courts, patent office officials, and inventors—about what patents were meant to protect. Everyone knew that manufactures and machines were at the core of the patent system.¹⁸ Agricultural and industrial machinery was almost synonymous with “patents.” For Thomas Jefferson and his cohorts, a piece of technology was readily identifiable: it had substance, and moving parts, and did something out in the practical world of farming or manufacturing. At the very least, for Jefferson, if you put technology in a bag and shook it, it would make some noise.

Against this background, it would have been seen as absurd for an entrepreneur to file a patent on a new finance technique such as publicly traded corporate shares, techniques for obtaining private financing for a bridge to compete with an existing bridge, or a security interest in uncut timber. These were the earmarks of commerce, of enterprise; laudable, surely, but something altogether distinct from the realm of “invention” and “the useful arts.”

Indeed, it might well have been argued that patents on such things were precluded by the British Statute of Monopolies—a statute which itself grew out of abuses in the grant of exclusive franchises in various lines of business such as trading cards, alehouses and various staple products.¹⁹ The Statute of Monopolies, after all, prohibited the grant of monopoly rights in various lines of business, rights which had been used as a device to raise public revenue and reward court favorites.²⁰ And the line that the Statute drew precisely reflected the eighteenth century view that technology was special: only new and useful inventions could receive patents, because these were the only property rights that could enhance social wel-

18 See, e.g., DAVID A. HOUNSHELL, *FROM THE AMERICAN SYSTEM TO MASS PRODUCTION, 1800-1932*, at 5 (1984) (describing the McCormick reaper and the Singer sewing machine, classic examples of nineteenth century technology).

19. See Robert P. Merges & Glenn H. Reynolds, *The Proper Scope of the Patent and Copyright Power* (Nov., 1998) (working paper on file with author). It has recently been argued that the report of one early case in this area may well have been distorted. See Jacob Corre, *The Argument, Decision, And Reports of Darcy v. Allen*, 45 EMORY L.J. 1261, 1266 (1996) (“[T]he opinion in *Darcy v. Allen* should not be viewed as a late-Tudor instance of the kind of explicit and concerted constitutional attack on the Crown that contributed so significantly to the Civil War forty years later.”).

20. See MERGES, *supra* note 7, at 6.

fare.²¹ The old, discredited monopolies in everyday household items were understood to contribute nothing new, except higher prices.²² A reasonable extension of the underlying logic of the Statute would have been to preclude patents on new business techniques, on the ground that the statute prohibited any special privilege to a commercial enterprise, as opposed to a new technology.

Alas, no such consensus on what patents are meant to protect exists today. Computer software—a sort of quasi-machine constituted out of written programs, or a written code that does machine-like work—has clouded and confused our working definition of “technology.” Program writers, or “software engineers,” labor to solve complex, demanding problems with standard toolkits, much as a determined inventor worked to design a new textile machine or seed drill in Jefferson’s day. Just because the end product of today’s engineering mind is manifested in a string of bits, it is no less a piece of “technology” than practical solutions of old, expressed in wood or steel.

Patent lawyers, paid to push the outer limits of what is protectable, have responded to the new technological realities with remarkable creativity. In the realm of financial instruments and Internet business concepts such as Priceline.com, the ubiquitous presence of computer technology permits inventors and their lawyers to characterize new businesses as essentially new combinations of hardware and software, and in some cases as new software packages per se. Once the Wall of Jericho holding back the forces of software patents was breached—and there can be no doubt anymore that the breach has occurred²³—the way was open for computer-related business concepts to be patented. When these software-embedded concepts are characterized as novel computer programs, there is little to separate them from any other computer program. They are therefore just as patentable as any other software.²⁴ QED.

21. *See id.*

22. *See* Statute of Monopolies, 1623, 21 Jam., ch. 3, § 6 (Eng.) (stating that no patents that raise the “prices of commodities at home, or hurt of trade, or generally inconvenient ... [will be allowed]”).

23. The surest sign that software is widely accepted as appropriate patentable subject matter is that we are beginning to see software-related patent infringement cases that do not even mention section 101 as an issue. *See, e.g.,* Enpat, Inc. v. Microsoft, Inc., 26 F. Supp. 2d 806 (E.D. Va. 1998).

24. Indeed, there is a fair argument that a business concept is patentable whether or not it is implemented on a computer:

[In *State Street Bank*] the Federal Circuit indicated that whether an invention is directed to patentable subject matter under § 101 does not

The acceptance of business concept patents is not due simply to the underlying technology. Another important cause is the shifting baseline in the intellectual property field. Beginning in the earliest days of the patent system, and extending until perhaps as late as the early 1980s, the legal system assumed that intellectual creations were *not* protectable *unless* (very) good cause was shown. Today, it often seems the opposite. We now ask: why not protect a new form of intellectual creation? We're protecting everything else like it.²⁵

This leads us back to the Constitution. At the practical level of this essay, we ask: does Article I clause 8 tell us what to do about the Walker Digitals of the world? Can we constitutionalize the implicit understanding of the framers and early patent system actors that patents are at their core about machines and manufactures—about nineteenth century technology, in other words? The question seems to answer itself. Given a constitutional provision rooted in a blind faith in “progress,” we cannot read in historically contingent limitations on patentable subject matter.²⁶ Put simply, there are no plausible subject matter limits, express or implied, in this broad, enabling clause.

If we want limits, we must look to Congress and the courts to provide them. While Congress is still a possibility, it has shown little inclination to limit intellectual property rights in recent years. And as for the courts, we have their definitive answer in such cases as *State Street Bank*, mentioned earlier.²⁷ In upholding claims to software inventions, the court has sup-

depend on whether a “physical” transformation takes place or whether the claim is directed to a process or a machine. From this, it might then follow that a claimed process for, e.g., performing the function similar to Signature’s invention, is patentable even absent its use with a computer. As long as the variables represent some set monetary values, it arguably should not matter who or what does the “transforming.” After all, regardless of the transforming mechanism (e.g., machine or human), the invention can be said to have “practical utility,” and produce a “useful, concrete and tangible result.” As can be appreciated, arguments can also be made as to why *State Street* might not be so broadly interpreted.

Alter, *supra* note 4, at 28 (citations omitted).

25. See generally Robert P. Merges, *The Economic Impact of Intellectual Property Rights: An Overview and Guide*, 19 J. CULTURAL ECON. 103 (1995).

26. For an argument that the phrase “for limited times,” in the historical context of the Intellectual Property Clause, does set limits on Congress’ ability to extend individual patents and copyrights through so-called “private bills,” see Merges and Reynolds, *supra* note 19.

27. See *supra* note 4.

plied a broad interpretation of the statutory classes of “process” and “machine.” Unless something changes in the statute or the courts’ interpretation of it, we can expect few subject matter limits on patents. This is one important reason we can expect to see continued increases in the number of patent applications, and hence, continued pressure on patent quality, in the coming years.

IV. EVALUATING THE PATENT EXAMINATION SYSTEM

At first glance, it ought to be easy to predict whether the whole idea of business concept patents makes any sense. All we need is a simple theory of when patents are necessary to call forth innovation, and when they are not. Once we have understood the category of things that will not be created in the absence of business method patents, we then ask the simple question: is it worth the social costs of granting exclusive property rights so that those things will be created? In essence, can we design a property right so that we gain more than we give up by granting such a right to those who qualify for it?

Probably not. Put simply, there is no easily-identified “ideal” menu of property rights for a given economy at a given moment in time. While it is clear in theory that only efficiency-enhancing property rights ought to be granted,²⁸ it is not always so simple in practice to tell what they are.

It is virtually impossible to determine—at least at this time—if truly valid business concept patents are a net drag on the economy, a net plus, or neutral. So I am *not* going to argue about that. But I *will* argue that we need to pay very close attention to the process by which these patents are granted, because, where the net effects are possibly negative, there is even more reason than usual to be concerned about improperly granted patents. I will therefore focus my attention on improving patent quality, generally.

My proposals are directed primarily at the PTO, the courts, and Congress. Because there is very little chance that any of these entities will act on them, I can be bold. My goal is to convince these people that while we may not be sure whether business concept patents are good or bad, we do know that *bad* business concept patents are bad. We must take steps to limit the damage from the ones already out there, and prevent more of them from issuing.

28. See DOUGLASS C. NORTH, INSTITUTIONS, INSTITUTIONAL CHANGE AND ECONOMIC PERFORMANCE 52 (1990).

Prevention can be achieved best by revamping the patent examination system in the PTO. Business concept patents are not the only reason to make these changes, but they are certainly a sufficient reason. And they might be just the straw that tips the balance in favor of much-needed reforms without which the proud tradition of the U.S. patent system is sure to continue its slow decay.

A. Why Is Patent Quality So Poor?

There are persistent reports that patents in the software area, and perhaps especially, patents for "business methods" implemented in software, are of extremely poor quality.²⁹ People familiar with the technology involved and the history of various developments in it report that patents in this area are routinely issued which overlook clearly anticipating prior art.³⁰ The average number of prior art references cited in software-implemented business concept patents has been said to be fewer than five.³¹ Three out of the five, on average, are citations to other U.S. patents, leaving an average of two non-patent citations per patent. What is disturbing about this figure is that patents have only recently become available for this technology. Consequently, we would expect that most of the prior art in this field would be of the non-patent variety. There is every reason to believe that there is a vast volume of non-patent prior art in the software-implemented business concept field, as is widely believed to be the case with software patents in general. Given that businesspeople have been pioneering new concepts since commerce began, and that Internet commerce has seen exponential growth in recent years,³² very few of the

29. See, e.g., Brenda Sandburg, *Patent Applications Flow Freely*, LEGAL TIMES, Feb. 22, 1999, at 12; Kenneth W. Dam, *Some Economic Considerations In The Intellectual Property Protection Of Software*, 24 J. LEGAL STUD. 321, 369-71 (1995) (discussing many of the problems with patent quality that had been identified with respect to software patents, and voicing optimism that problems can be addressed).

30. See Andrew M. Riddles & Brenda Pomerance, *Software Patentee Must Conduct Own Search: Prior-Art Searches Made By The Patent Office Often Are Not Thorough Enough To Be Trusted*, NAT'L L.J., Jan. 26, 1998, at C19. (accusing PTO of being little better than a "registration process" for some kinds of software patents).

31. See Greg Aharonian, *17,500 software patents to issue in 1998*, INTERNET PATENT NEWS SERVICE (Oct. 18, 1998), available at <<http://lpf.ai.mit.edu/Patents/ipns/ipns-19981018.txt>>.

32. See, e.g., *Quantel, Ltd. v. Adobe Systems, Inc.*, 1997 U.S. Dist. LEXIS 16779 at *14-23 (D. Del. Sept. 22, 1997) (jury verdict invalidating software patents; special verdict form shows numerous "prior public use" references). Cf. MERGES, *supra* note 7, at

developments in this area have found their way into patents. They are reflected instead in actual businesses, business plans, the financial services industry literature, and the like.³³ It therefore seems likely that many of the patents being issued in this area overlook highly relevant prior art. Thus, the error rate for these patents is likely to be quite high.³⁴

No doubt part of the problem is that the patent system has only recently begun to issue patents in this field. Thus, perhaps we can expect some low quality patents now, until the patent system has time to adjust. This has certainly been our experience in other fields: there were numerous complaints in the early years of biotechnology and software patents that the PTO was allowing too many overly broad patents.³⁵

At the same time, the scope of the problem seems to be worse this time. Partly, this is a simple matter of overall volume: the PTO has experienced a very rapid increase in the number of patent applications filed in the past few years.³⁶ For reasons that will be explained later,³⁷ there are numerous incentives inside the PTO to issue rather than reject patent applications. As a consequence, the number of patents issued has also grown sharply in the past few years.³⁸

416; George Gates, *Trade Secret Software: Is It Prior Art?*, 6 COMPUTER LAW. 11 (1989).

33. See OFFICE OF TECHNOLOGY ASSESSMENT, U.S. CONG., FINDING A BALANCE: COMPUTER SOFTWARE, INTELLECTUAL PROPERTY AND THE CHALLENGE OF TECHNOLOGICAL CHANGE 24 (1992) (noting need to "fill[] in" the prior art to improve software patent quality).

34. The Attorney-Advisor to Commissioner Christine A. Varney of the Federal Trade Commission has this to say on the topic:

Given continuing data and expertise problems, any expansion of the scope of statutory subject matter will inevitably result in the issuance of more patents that do not meet the statutory requirements of novelty and nonobviousness, but instead have the potential to block further software development. Given the non-public nature of the patent application process, the absence of effective post-award review and the substantial transaction costs associated with defending patent infringement litigation, many improvidently granted patents are likely to go unchallenged.

J. Beckwith Burr, *Competition Policy And Intellectual Property In The Information Age*, 41 VILL. L. REV. 193, 204 (1996).

35. See, e.g., Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 905-06 (criticizing an early monoclonal antibody diagnostic kit patent).

36. See fig.1 *infra* p. 601.

37. See *infra*, Part V.A.

38. See fig.1 *infra* p. 601.

The concerns about quality, especially in light of the data on overall volume, point to one conclusion: the patent system is in crisis. Therefore, this is an opportune moment to take a step back and ask an important preliminary question: how thorough should patent examinations be? A thorough analysis of these fundamental issues will help immensely in deciding whether the reported crisis is genuine.

B. Sketching an Ideal Patent Office

In this section, I discuss how we would design a patent system if we were starting from scratch today. It seems to me there are four important subissues here: (1) How much time and effort should the PTO spend on each patent application, and is there any way for the PTO to sort patent applications by anticipated economic value? (2) Inasmuch as the competitors of a company that receives a patent will potentially bear the costs of an improperly granted patent, is there any way to harness their self-interest and their intimate knowledge of the technology to bear on the patent application process? (3) What is the proper “division of labor” between the PTO which issues patents and the courts which later review them—in essence, what is the ideal standard of review for the validity of an issued patent? And (4) what are the optimal remedies and punishments for acquiring an invalid patent or asserting it against competitors?

I will address the first two issues in this article, leaving the others for later analysis. Of course, it should be understood that changes in one area of the system may have important consequences for other areas. But we must start somewhere, and the PTO—as the government agency that serves as the first and important line of defense against socially wasteful patents—is as good a place as any.

1. *The function of a patent office*

It is curious that in all the vast economic literature on patents, virtually nothing has been written about the functioning of a patent office. When patent granting authorities are mentioned, it is usually as a “black box” bureaucracy out of which patents emerge.³⁹ There is scant literature on auctioning research projects.⁴⁰ This literature describes a proto-patent office that auctions off the right to investigate and develop a discrete and identifiable technological “opportunity.” But it seems self-evident that this is farfetched enough to be disregarded. Technology is rarely so readily

39. Cf. Merges, *supra* note 25.

40. See, e.g., Yoram Barzel, *Optimal Timing of Innovations*, 50 REV. ECON. & STATISTICS 348 (1968).

identifiable, and its future prospects rarely if ever well enough understood, to induce reasonable bidding. In addition, the potential developers of a technological prospect are unlikely to be so readily identifiable as permit them to assemble in a single auction.

Because of the dearth of antecedents in this area, we will have to proceed on first principles. In that spirit, I offer the following description of the goals of an ideal patent system. This is then tempered with some real-world considerations. But first, some first principles.

2. *The goals of a patent office*

On one level, of course, it is easy to describe the goal of the PTO. It should follow its statutory mandate closely, issuing only patents that its enabling legislation permits or deems desirable. According to this view, no patent which lacks statutory novelty, which is obvious in light of the prior art, or which includes claims that are not enabled under the terms of the statute, ought ever be issued.

This general statement cannot be faulted on one level: the issuance of an invalid patent results in some social costs that could have been avoided. (All patents, even those that are in fact completely valid, involve social costs; the only issue here is whether those costs could have been avoided by more thoroughly searching the prior art to find invalidating references.)⁴¹ Of course, the costs of invalid patents include the direct costs of filing and prosecution. There is also a myriad of indirect costs, including: unnecessary licensing fees; foregone research opportunities, abandoned or avoided by the patentee's competitors who fear infringement liability; and the activities of rent-seekers⁴² who may respond to the combination of lax patent standards and robust rewards to patentees by diverting excessive resources out of productive activities and into the "patent game."

41. Implicit in this statement is that the technology at issue in the patent would be disclosed and/or commercialized even if no patent were granted. Put another way, the social cost is avoidable but the benefit is still realized. This is a bedrock assumption of our patent system. Our rules of novelty and nonobviousness assume that if technology is available "off the shelf" then someone will implement it without the need for any special property right. See MERGES, *supra* note 7, at 259-63 ("Novelty and the Economics of 'Search'").

42. The term "rent-seeker" refers to those who seek a supra-competitive return. The usual sense is negative; thus, one who seeks such a return from an illegitimate, non-welfare-enhancing source is a rent-seeker. An example is a person who makes campaign expenditures on candidates who promise to back legislative action that profits one or few at the expense of the many.

The fundamental assumption behind public expenditures on a patent office in the first place is that, as a society, we do not want to bear the costs of a significant number of invalid patents. Indeed, as described below, historically the current system of a professional corps of patent examiners grew out of our disastrous experience with a patent registration system run amuck. The social costs of large numbers of invalid patents were considered high enough to justify the significant expense of setting up a real patent office.

But does this necessarily translate into a goal of zero invalid patents? If this is the goal, then the issuance of a single invalid patent—one that is in fact anticipated, or obvious, but that the patent office has erroneously issued—means the office has failed.

This is not only unrealistic, as we will see below; it is also inconsistent with certain signals we receive from our patent statute. If no invalid patents are supposed to be issued, then why have the independent court review of patent validity called for by our statute? Why have a mere “presumption” of validity rather than a “conclusive presumption,” i.e., an unreviewable determination that patents, once issued, are valid for all time? Perhaps one reason is that, sometimes, prior art does not mature or come to light until after patents are issued. But this could be addressed by phasing in a conclusive presumption after some period of time, much as a trademark can in some cases become “incontestable” five years after it is first registered.⁴³ An argument against instituting this kind of delayed presumption is that, given the high social cost of an invalid patent, even prior art discovered very late in a patent’s term should be brought to light.⁴⁴ The

43. See Lanham Act § 33(b), 15 U.S.C. § 1115(b) (1998).

44. In an interesting treatment of related issues published as this article went to press, Ian Ayres and Paul Klemperer work the other side of this issue. See Ian Ayres & Paul Klemperer, *Limiting Patentees’ Market Power Without Reducing Innovation Incentives: The Perverse Benefits Of Uncertainty And Non-Injunctive Remedies*, 97 MICH. L. REV. 985 (1999). They study the relative benefits of uncertainty and delay in the enforcement of patents. Their overall point is that uncertainty and delay in enforcing patents can at times permit limited entry that erodes the patentee’s ability to price at the monopoly level. Their basic insight is that even a small amount of uncertainty regarding patent enforceability can have significant positive effects on social welfare, with much more limited negative consequences for the patentee’s incentives. (This flows from the fact that, near the monopoly price under conventional assumptions, price increases benefit the patentee only a small amount while producing very significant dead weight losses to consumers.) The authors consider a variety of doctrines that might be enlisted to increase *ex ante* uncertainty, including patent standards. This leads them to argue in favor of “underinclusive” patent standards, i.e., those that might permit more invalid patents to survive longer, and of relatively lax patent review (at the margin) by the PTO. See *id.* at

contrast with trademarks is obvious; at least traditionally, these are much weaker rights. Further, because there are many words and logos that might serve as adequate substitutes for a trademarked term, the social cost may be small enough that incontestability is a valid protection.

3. *Pros and cons of a simple registration system*

That Congress has chosen not to rely exclusively on administrative determinations of validity tells us something about the proper role of our patent system, something that we can build on in thinking about how to reform that system. But before we go on, we might want to consider the opposite extreme: why not revert to a registration system, similar to the one that was in effect between 1793 and 1836?⁴⁵ Why not, in other words, shift *all* the burden to the private sector, by registering any patent that comes along and letting the parties sort things out in litigation?

The argument in favor of registration is easy to make; it is what justifies the current copyright registration system. There are many copyrighted works that have either low intrinsic value (a brief trade press article, a schlocky picture), or have many close substitutes (most songs, many “genre” novels such as romances or mysteries), or both. To spend governmental resources sorting the good from the bad would be a waste of time. Instead, the copyright system lets private parties choose which copyrighted works are valuable enough to examine in detail. Then, in the course of litigation, the parties who deem it worthwhile will spend money describing why the copyrighted work is or is not protectable. Private sorting is more efficient.

This system was tried and rejected for patents, largely because of the high social cost. Private industry and Congress both concluded that the high cost of registering invalid patents was not worth whatever benefits

1025-26 (“On the margin, [their argument] militates against statutory or regulatory rules that are known *ex ante* and instead militates in favor of common law standards that often produce relatively delayed and uncertain adjudication—particularly if the common law is underinclusive.”). While their general argument is intriguing, it is submitted that uncertain PTO review is a poor way to implement it. The third-party costs—in the form of researching prior art, obtaining patent opinion letters, and revising research plans to avoid masses of uncertain but potentially valid patents—are simply too high. Best to apply their insights in other areas they explore, such as the doctrine of equivalents and the standard for granting preliminary injunctions.

45. For an excellent account of the problems with the patent registration system, and a thorough discussion of the genesis of the 1836 Act, see Edward C. Walterscheid, *The Winged Gudgeon—An Early Patent Controversy*, 79 J. PAT & TRADEMARK OFF. SOC’Y 533 (1997).

were provided by this low “entry barrier” to inventors.⁴⁶ Even in the early nineteenth century patent litigation was complex, and therefore expensive.⁴⁷ Also, because courts then (and now) are not necessarily well trained in technology issues, the risk of error at trial was significant. Hence, there was even more expense, in the form of appeals to have an erroneous trial results reversed. Because of the cost and the potential for error, the threat value of even an invalid patent was substantial.⁴⁸ There were assertions that patents were being used to hold up bona fide manufacturers.⁴⁹ As patents became associated with rent-seeking rather than innovation, the net result was to undermine the integrity of the patent system as a whole, and thereby (presumably) reduce the incentive to innovate that patents are supposed to represent. Innovation was replaced with rent-seeking, as unscrupulous people and firms played the game of patent extortion.⁵⁰ Thus, patent examination—an increased public expenditure on patent quality—was instituted on a formal, regularized basis.

An economist reading this history would conclude that it is a classic illustration of government intervening to overcome externalities.⁵¹ Private parties, responding only to market signals, produced too many invalid patents; patent litigation mushroomed; and (again, presumably) there was an overall negative effect on innovation.⁵² In stepped the government, after

46. *See id.* at 535-36.

47. *See* STUART BRUCHEY, *ENTERPRISE: THE DYNAMIC ECONOMY OF A FREE PEOPLE* 230 (1990) (describing Eli Whitney’s frustration at long, expensive and “fruitless” litigation over the cotton gin patent).

48. *See* Walterscheid, *supra* note 45, at 548.

49. *See id.* at 549 (quoting a federal judge, who declared “[The] very great and alarming facility with which patents are procured [under the registration system] is producing evils of great magnitude. It encourages the flagitious peculations of imposters, and the arrogant pretensions of vain and fraudulent projectors ... the community suffers under the many diversified extortions”).

50. *See id.*

51. *See, e.g.*, HAL R. VARIAN, *INTERMEDIATE MICROECONOMICS: A MODERN APPROACH* 545-65 (3d ed. 1993).

52. One commentator wrote:

[T]he major defect of the Patent Act of 1793, which remained the law of the land until 1836, [was] ... that anyone could obtain a patent for anything, merely by paying the requisite fee and meeting the ministerial requirements imposed. It mattered not that the supposed invention had already been patented or had long been known and used. The threat of litigation was sufficient for the owners of apparently invalid patents to obtain substantial royalties from literally hundreds and thousands of farmers, small businessmen, and artisans for whom it truly was cheaper to pay than to be involved in expensive and perhaps ruinous litigation.

having determined that the expenditure on patent examination would increase the net benefits of the patent system by reducing the social cost of an excessive number of invalid patents.

So, from the history and structure of our current system we learn these lessons: neither pure public (administrative) proceedings, nor pure private (registration system) proceedings are efficient. Our patent system envisions a mixture of public *and* private expenditures to determine the validity of patents. Indeed, it is part of a larger theme in patent law: the division of labor between the public and private sectors in the issuance and enforcement of these property rights.⁵³

C. Optimal Public Expenditures on Patents

Before we can determine the ideal mix of public and private expenditures on patents, we must address the public side of the ledger in isolation. The question of interest here is how to determine the correct magnitude of public expenditure on patent quality control.

In theory, the answer is simple. Following conventional principles, the public expenditure should increase until it is not worth increasing it any more—until the marginal cost and benefit are equal. To determine this, we only need to know: (1) the cost to the patent office of each additional unit of validity information; and (2) the estimated social cost of each patent, expressed as a function of the volume of validity information processed for that patent. This second element is needed to reflect some sense of *reliability*: if we know that an invalid patent on average costs society \$X, and we know that each additional unit of search effort reduces the probability that the patent office will issue such an invalid patent by Y%, then we can determine the expected savings to society resulting from a more thorough and careful search of the prior art.

1. *The benefits of sorting applications*

It follows from the preceding that patent applications should be subject to differing levels of scrutiny depending on how much social cost they entail. Applications for patents that would be very costly to society—because they are very broad, for example, or because there are no good sub-

See Walterscheid, *supra* note 45, at 533.

53. The PTO only issues patents, it does not identify, locate, or sue infringers. At the same time, a private party cannot bring an enforcement action until his or her patent is granted, and even then this private action may be “stayed” while the patent is undergoing reexamination.

stitutes for the patented technology—ought to be examined more closely than those for minor improvements, gadgets, or novelties.

If we express this in simple qualitative terms, it means that, ideally, we would sort patents according to their prospective social cost. We could then allocate the available search resources so as to spend more resources on the patent applications likely to mature into patents with a high social cost, and less on the on the applications likely to produce patents with little social cost.

2. *The problems (and politics) of sorting*

While it is theoretically possible for the PTO to perform such a sort, there are currently significant barriers to doing so. One is informational: it is difficult at the time of filing to determine which applications may mature into high social cost patents. Patents are usually filed early in the development phase, and the inventor often has little idea whether or not the technology will “pan out.” Hence it makes sense to delay the sorting for as long as possible. (In Japan, and to a very limited extent in the United States, applicants themselves can in some cases approximate this: they can either “activate” a pending application through filings with the patent office, or leave it dormant.⁵⁴ Early activation might be taken in this context as a proxy for higher expected private value, and hence higher social cost.)

But there is another barrier to sorting. The history and culture of our patent system reflects a broad egalitarian streak. In the patent system, by custom “all patents are created equal.”⁵⁵ Any mechanism for separating patent applications would necessarily buck this tradition. It could of course be argued that proportional rationing of scarce examination resources still meets the test of equal treatment (in the sense that similar applications would be treated similarly). And perhaps this would prove persuasive. But there is still the possibility that any effort to segregate patents

54. See Hideo Kodama & Jeffrey D. Tekanic, *Reducing the Costs of Obtaining and Maintaining Japanese Patents*, 81 J. PAT. & TRADEMARK OFF. SOC'Y 117, 127 (1999) (describing 7 year deadline to request examination). See also 35 U.S.C. § 111(b) (describing provisional patent applications). Provisional patent applications are not examined by the PTO and can be replaced at any time up to one year from filing with a normal patent application. See *id.* This effectively allows a one-year “option” period for inventors to delay examination of a patent application.

55. *The Public Interest and Private Patent Bills: Senate Hearings on Patent Extensions (Private Patent Bills)*, 102d Cong. 102-824 (1991) (statement of Prof. Robert Merges). An alternative would be to provide a “second tier” of patent protection to less significant inventions, as is done in foreign “utility model” protection schemes. See generally Mark D. Janis, *Second Tier Patent Protection*, 40 HARV. INT'L L.J. 151 (1999)

into various classes would be perceived with hostility by patent traditionalists. In this setting, it may prove very difficult to obtain approval for any effective sorting mechanism, which would by definition deviate from strict equal treatment.

3. *Second-best solutions*

If sorting is impossible for political reasons, what else might be done? Two things: (1) raise the standard of patentability and/or the filing fees, in order to induce applicants to sort out the least potentially valuable investments on their own; and (2) make a rational guesstimate regarding a reasonable *average* expenditure on examination, and set the overall patent budget accordingly.

The first proposal raises the cost of applying for a patent. In marginal cases, where the probability of receiving a patent is low, the value of the invention low, and the cost of applying for the patent high, prospective applicants will choose not to file. The filing fee might make the most sense as a screen; it could potentially raise revenue, and a fee increase is much easier to implement than increasing the standard of patentability. The easiest way to raise standards, conceptually, is to tighten the nonobviousness requirement of section 103. However, this is a notoriously subjective standard, and it may prove difficult, not only to draft a tightened requirement, but also to make it stick.

The second proposal is perhaps more workable: all inventors would presumably benefit from a rationally derived PTO budget.⁵⁶ In theory, the approach would simply be to set the PTO budget equal to the total social cost of all invalid patents. Then, assuming equal expenditure on each patent application, the PTO would spend an amount equal to the average cost of an invalid patent.

Note that, while valid patents would survive the examination process, so too would a certain number of invalid patents. These would be those patents that cannot be cost-effectively eliminated at the examination stage. Such invalid patents have a close corollary in the economic literature on tort law: accidents that cannot be avoided at reasonable cost.⁵⁷ As with these accidents, invalid patents that are too expensive to weed out must be tolerated. By definition, the money that would be spent to eliminate them is better spent elsewhere.

56. Currently, the budget is largely a function of the fees the office collects, minus some money that Congress skims off for the general fisc. See 1995 U.S. PAT. & TRADEMARK OFF. ANN. REP. 47.

57. See ROBERT COOTER AND THOMAS ULEN, LAW AND ECONOMICS 326-71 (1988).

a) Why not penalize holders of invalid patents?

The tort/accident analogy suggests an interesting question: why does the legal system fail to require a patentee whose patent is invalidated to compensate an alleged infringer, all competitors, or even society in general (via a fine)? In tort law, legal damages are the negative incentive that induces precaution on the part of a potential tortfeasor. In our discussion so far, we have described the PTO as the relevant agent to determine the appropriate level of "precaution" against invalid patents. Why not shift some, or all, of this cost to the applicant?

The answer must be that we are concerned that such a rule would deter too many patent applications, and hence too much valuable inventive activity. Consider that, even though a good deal of the prior art that can invalidate a patent is publicly available, much is not. Internal developments at a competitor firm can manifest themselves in a number of types of prior art, and there is usually no way for a patent applicant to find out about this activity until after—sometimes, well after—a patent application is filed. If no amount of pre-filing search could have turned up this evidence, it is harsh and inefficient to punish a patent applicant when it comes to light.

On the other hand, where an applicant *did* know about a piece of relevant prior art, and failed to call it to the attention of the examiner, Rule 56 of the PTO practices results in the invalidation of the patent.⁵⁸ In addition, in extreme cases a patent applicant can be liable for up to treble damages in antitrust if he or she knowingly prosecuted an invalid patent application with an eye toward monopolizing a product market.⁵⁹ This is rare, however.

b) Who is the cheapest cost-avoider?

To complete the patent/tort analogy, it is appropriate to ask who is the cheapest cost avoider. In tort law, this consideration answers questions such as who, between two parties, ought to bear liability if there is an accident; and how should that liability be apportioned, if at all?⁶⁰

By analogy, we might ask: who is in the best position to avoid the social costs of an invalid patent? One possible choice for the cheapest cost-

58. See 37 C.F.R. § 1.56 (1999). See also Harry F. Manbeck, Jr., *The Evolution and Issue of New Rule 56*, 20 AM. INTELL. PROP. L. ASS'N Q.J. 136 (1992).

59. See, e.g., *Walker Process Equipment, Inc. v. Food Machinery & Chemical Corp.*, 382 U.S. 172 (1965); *Handgards, Inc. v. Ethicon, Inc.*, 743 F.2d 1282 (9th Cir. 1984), cert. denied, 469 U.S. 1190 (1985).

60. The classic reference here is GUIDO CALABRESI, *THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS* (1970).

avoider is a public patent authority. The reasons for a public patent authority—both rational and political—have been sketched in the above sections. They boil down to these:

- Up to a certain point, there may be economies of scale in doing “commodity” prior art searches such as searches of widely-available scientific and technical articles and prior patents;
- There is value in a public examination function which guarantees some minimum quality level to patents, in part to prevent the most egregious patent “strike suits” or extortion attempts that depend for success on the high cost of patent litigation; and
- It is politically desirable to shift some of the costs of patent searches from small inventors to the patent office.

As we shall see, much of the information that bears on patent validity is held by private parties, and especially by the patent applicant’s competitors. This leads me, in a later section, to champion an opposition system.⁶¹ Such a system would get more of this information into the patent examination system, and would do so at an earlier date than under the current system. Before we get there, however, we must complete our discussion of the ideal role of the public patent authority.

D. Examining the PTO

I have to this point laid out the case that the patent system is in crisis. And I have hinted that part of the answer should come in the form of increased private investment in patent quality—in the form of an opposition system. But another part of the answer is on the public side of the ledger. Thus, we turn now to reforming the PTO.

1. The PTO’s examination budget

We begin our discussion of PTO reform by looking at the basics of our examination system. The PTO in its modern form was put in place in 1836.⁶² Before then, except for a brief “heroic” period when Thomas Jefferson and others administered it, inventors merely registered patents.⁶³ Validity was determined solely in district court litigation. The advent of a

61. See *infra* Part V.C.

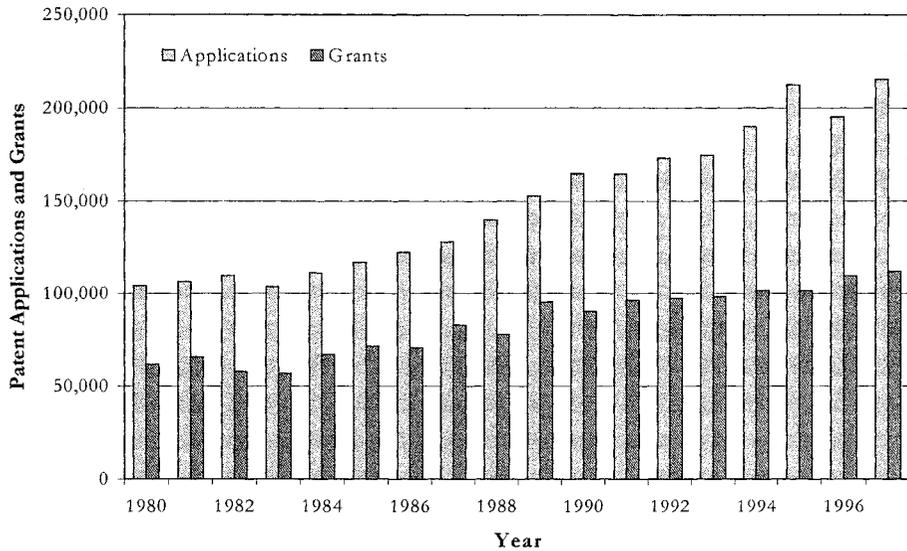
62. See MERGES, *supra* note 7, at 9-10.

63. See *id.*; Walterscheid, *supra* note 45, at 534.

modern examination system initiated the current mix of public and private review of patent validity.

Today the PTO is a large institution. Figure 1 shows the general trend in patent applications and grants over time:

Figure 1. U.S. Patent Applications and Grants, 1980 – 1997⁶⁴



The fees that inventors pay for applications, issuance, and renewals now exceed \$674 million per year.⁶⁵ The office even generates a surplus, which Congress routinely seizes for the general fisc.⁶⁶ The revenue picture has been changing drastically in recent years; from 1990 to 1991 the patent processing fees collected nearly doubled, from \$175 million to \$290 million. By 1993, the number had jumped to \$423 million.⁶⁷

The PTO spends this significant amount of money on a number of things, including policy development, international coordination, and, of course, patent examination.⁶⁸ The latter category includes not only initial examinations, but also interference proceedings to determine priority amongst rival claimants, reexaminations, reissues, and a number of related

64. See 1997 U.S. PAT. & TRADEMARK OFF. ANN. REP. 85 tbl.2; *id.* at 87 tbl.6.

65. See *id.* at 73.

66. See, e.g., *id.* at 35.

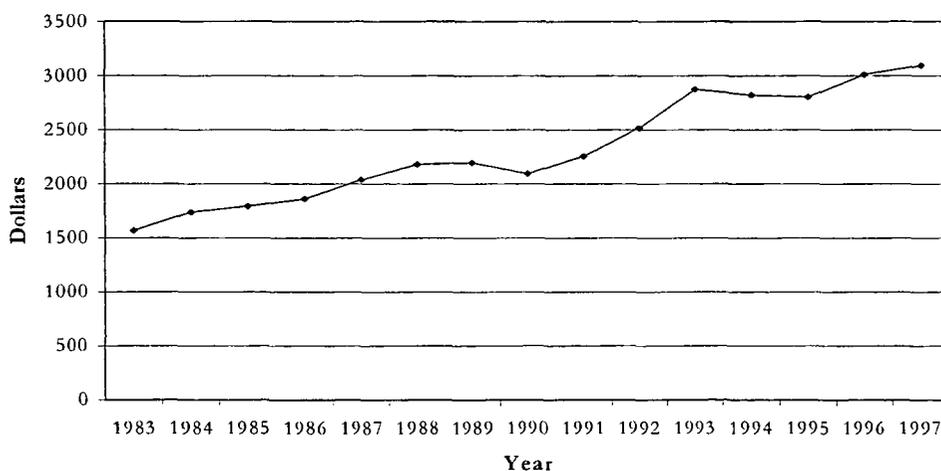
67. See 1993 U.S. PAT. & TRADEMARK OFF. ANN. REP. 51.

68. See generally 1997 U.S. PAT. & TRADEMARK OFF. ANN. REP.; 1996 U.S. PAT. & TRADEMARK OFF. ANN. REP.; 1995 U.S. PAT. & TRADEMARK OFF. ANN. REP.

activities. For brevity, I will refer to all of these as "examination expenses."

In recent years, the PTO has received on the order of 230,000 patent applications each year.⁶⁹ Given current revenue, that means that it has available, in theory anyway, approximately \$3000 per patent. Figure 2 shows the trend in estimated expenditure per patent over time. On one level these are reassuring figures. It is now a truism that intellectual property is the key asset in the emerging economy. Patents are obviously an important component. Thus, it arguably makes sense that we as a society have increased our spending on the examination of patents. Patents are potentially worth more than they were in the past; thus, the cost of an improperly granted patent might also be presumed to have risen.

Figure 2. PTO Funding Availability Per Patent Application, 1983 - 1997⁷⁰



69. See 1997 U.S. PAT. & TRADEMARK OFF. ANN. REP. 85 tbl.2.

70. See 1997 U.S. PAT. & TRADEMARK OFF. ANN. REP. 41 (showing the PTO's revenues for fiscal years 1996 and 1997); *id.* at 85 tbl.2 (showing the number of patent applications filed during 1996 and 1997); 1995 U.S. PAT. & TRADEMARK OFF. ANN. REP. 52 (showing the PTO's total resources for fiscal years 1994 and 1995); *id.* at 87 tbl.1 (showing the number of patent applications filed with the PTO during fiscal years 1994 and 1995); 1993 U.S. PAT. & TRADEMARK OFF. ANN. REP. 49 tbl.2 (showing the total funding PTO funding availability, per fiscal year, from 1983-1993); *id.* at 54 tbl.6 (showing the number of patent applications filed with the PTO each fiscal year, from 1983-1993). Because it is not always possible to ascertain the *patent* budget, alone, from the PTO's annual reports, I have substituted total funding availability. This is possible because we are merely observing the ratio of funding to patent applications. Note that

2. *Setting the ideal PTO budget*

The growth in the PTO budget and expenditures raises an obvious question: is this enough in some absolute sense to do a good job? Should we be increasing expenditures even faster, keeping them constant, or perhaps reducing them?

The economic literature on property rights provides some guidance here. For one thing, it shows that at some point the potential value of an asset is high enough to justify establishing or strengthening property rights over it.⁷¹ It is implicit in academic work along these lines that a new system of rights requires a new administrative infrastructure: land registries, title recording procedures, and the like.⁷² The simple notion is that, given the economic advantages of stronger property rights, at some point in economic development the extra public expenditure on additional property rights infrastructure creates a net benefit.

Beyond this simple statement, however, there is little guidance. Looking backward, we can see that it made sense to institute a title registry system, or to clarify the law of mining claims. But we do not have the tools to determine *in advance* the ideal public expenditure level for any given property right. In the patent context, it is entirely possible that the current budget is the right one, or at least a workable one. (I discuss some internal reforms to increase the productivity of these public expenditures below.) Given (1) the lack of information when patents are filed, (2) the fact that most technologies will not be economically viable or commercially successful, and (3) the high cost of separating out the potentially valuable inventions, it may make sense to continue to spend roughly what we do now on patent examinations.

At the same time, given the large increase in the private value of patents since the early 1980s, it is also plausible that the government should be spending more on examining patent applications. For as the average private value of a patent has increased, so has the social cost of an invalid patent.

the numbers do not change greatly when we substitute patent funding, alone, which is given in the 1997 U.S. PATENT & TRADEMARK OFFICE ANNUAL REPORT: the amount available per patent changes little, from \$3100 to \$2800. *See id.* at 73 (showing fee collections by category).

71. *See* Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. 347, 347-59 (1967).

72. *Cf.* Robert Ellickson, *Property in Land*, 102 YALE L.J. 1315, 1329-30 (1993) (describing origin of deed registries).

In an ideal world, the PTO would have a completely accurate prediction of the potential future value of a patent application. That is, the patent examiner would know: (a) the future rent stream that will flow from the patent if it is granted; (b) the number and value of future inventions that the patent application will spawn; and (c) competitive conditions in the market into which the invention will be sold, including alternative technologies and their cost.

The PTO could use this information to determine how much money to spend on the examination of each application.⁷³ The idea here is quite simple. The PTO would do a straightforward cost-benefit analysis. Following the well-known literature on the incremental value of information, the PTO would determine the marginal cost of each additional unit of patent examination effort.⁷⁴ The office would then calibrate this to the expected benefit from the patent application. The idea would be to tailor each patent examination to the potential future value of each patent application.

Notice that underlying this view of the PTO is the notion that it is the agent for all competitors and consumers who will be affected by the issuance of a patent. Under this view, the PTO has a simple job: to maximize social welfare by scrutinizing patents and allowing only those that survive a cost benefit-adjusted search process to issue.

Of course, this ideal world assumes that the PTO search and examination process is the most efficient one available.⁷⁵ What if an outside party has better information about patentability characteristics of the invention? Under these circumstances, it would be wise to permit the PTO to subcontract patent search and examination procedures to outside firms that

73. This is very similar to Posner's discussion of optimal filing fees in civil litigation. See RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 578-82 (4th ed. 1992).

74. See JACK HIRSHLEIFER AND JOHN G. RILEY, *THE ANALYTICS OF INFORMATION AND UNCERTAINTY* 180 ff. (1992). To rigorously pursue the statement in the text, one would need to specify some additional variables, most importantly the PTO's *a priori* probability assessments of the validity of the patent (based on its internal search results information, e.g., that a search of certain stringency leads to an identical finding, and thus lack of patentable novelty, in 20% of all cases). In search theory, new information—the product of the search—operates to modify earlier probability assessments. For an application of this Bayesian approach to patent validity, see Robert P. Merges, *Uncertainty and the Standard of Patentability*, 7 *HIGH TECH. L.J.* 1 (1993).

75. That is, the cost benefit analysis that the PTO conducts is strictly a function of its own internal costs of examination and search. In that case, the PTO's cost benefit analysis will not reflect the true social welfare calculus, but instead only a "local" cost benefit analysis.

have better information, better search technologies or that simply work more efficiently. These firms would be, in effect, "private patent offices." This would yield a better decision regarding the optimal expenditure on search and examination for each application.

If we push the notion of subcontracting a bit further, we arrive at an important policy recommendation. To some extent, the ideal outside search firm—the one with the lowest cost of acquiring relevant information—would be a firm with access to all the information available to firms that operates in the same industry as the patent applicant. Indeed, because at least some of this information is considered a trade secret, the truly ideal search firm is an *actual* competitor of the patent applicant. Fortunately, this notion of subcontracting search to competitor firms does not require a radical restructuring of the patent system. It already exists, in the guise of patent oppositions, which are available in Europe and Japan, and have been proposed for the U.S.⁷⁶

Even the simple analysis of information costs presented here is only a first cut. It surely would require modification. For example, search and examination are not the only functions performed by the PTO. Thus, the PTO's total search and examination budget must be weighed against its expenditures for such other functions as policy-making, international negotiations, legislative research, and general manpower and management issues. For this reason, it might make sense to put a cap on the total examination budget for the patent office. Unless we were willing to adjust patent application fees to make each patent applicant bear the precise cost of the search and examination for his or her patent application (which is too difficult and expensive to calculate), there would likely be some hard budget constraint that would be taken into account in the total search and examination budget. Even assuming a "simple optimization" view of the problem, prior art searching is likely to be subject to steep diminishing returns at some point. If the vast amount of benefit is obtained with the first few increments of search and examination effort, then a patent that is predicted to have very high value might be subject to search and examination which, at the margin, yields quite small benefits. In some absolute sense, taking the hard budget constraint just mentioned into account, extensive searches and examinations might not be considered a wise invest-

76. For an overview of these proposals, see Allan M. Soobert, *Breaking New Grounds in Administrative Revocation of U.S. Patents: Proposition for Opposition—and Beyond*, 14 SANTA CLARA COMPUTER & HIGH TECH. L.J. 63, 128-44 (1998). We pick up the argument for oppositions, and the related argument in favor at least of reforming the U.S. reexamination system, later in Part V.C.

ment on the part of the PTO or the public it represents. Put another way, a rough judgment about the marginal value of additional searching might be substituted for a more rigorous marginal benefit analysis.

So, we can summarize the discussion as follows. In an ideal world, the PTO would calibrate its search and examination to each individual patent application. Barring this, it could attempt some sort of primitive triage, separating trivial patents (e.g., for gadgets) from those with industrial promise, and the latter into “potentially significant” and “probably minor.” If the administrative costs prove too high, the idea of triage could be abandoned, and the PTO’s emphasis could return to determining a rational amount of money to spend on each patent, taking into account the value of the average patent and some rough sense of the social cost of granting invalid patents.

V. SOME SIMPLE SUGGESTIONS FOR IMPROVING THE EXAMINATION PROCESS

Based on what we have discussed so far, from a “division of labor” perspective, how would we state the goals of our patent system? Here is one attempt:

- Issue patents whose *average* validity rate meets social welfare objectives; and
- Disseminate information about issued patents, and structure procedures, to enable efficient private-party validity review.

With these straightforward goals in mind, and admitting that we cannot determine the ideal expenditure on patent quality, we turn to some simple suggestions for improving the productivity of those funds we decide will constitute the public investment in patent quality (i.e., the PTO budget).

A. Job Design

A recurring theme in the assessment of PTO performance is poor examination quality due to high examiner turnover. This boils down to two specific problems: (1) too few senior examiners; and (2) inadequate training for the revolving cast of inexperienced examiners.

The answer to the first problem is as simple as it is difficult to achieve: higher salaries for senior examiners. Until the PTO can make it more attractive to stay than to leave, people will continue to leave. One interesting point to consider is a radically higher salary structure for the most senior

examiners. If their productivity is high enough, it may well be worth it. The current salary structure is difficult to document, but it appears that both the absolute pay levels and the rates of pay increases lag behind equivalent measures in the private sector.⁷⁷ The increased expenditure on higher salaries for senior industrial researchers is apparently worthwhile, because we observe that it occurs in a wide variety of industries.⁷⁸ In theory at least, productivity goes up enough with seniority to make it worthwhile to pay much more. The same is likely true among patent examiners.

But raising salaries for senior examiners is not the only way to tackle the problem. The second problem could also be addressed by shifting expenditures to training for the most junior people. Currently, junior examiners complain that they receive very little effective training. There are official programs on the books, but they do not do much, according to junior examiners.⁷⁹ This is because the most effective trainers—the senior examining corps—do not have any incentive to spend any time training. The patent compensation system, a combination of base salary and bonus, directs their effort heavily toward their own examining activities. Bonus points are accumulated only for “dispositions,” i.e., final allowances or rejections of patents. Because of the nature of prosecution procedure, “final” rejections do not in fact always result in the end of the examination; post-“final” action amendments and the like are often permitted. Consequently, the only way to earn bonus points with confidence is to allow a patent application.⁸⁰ In any event, there are no bonus points for training younger examiners.

77. The entry level job descriptions for patent examiners list salaries ranging from \$20,588 to roughly \$60,000. See U.S. Pat. & Trademark Off., *Patent Examiner Recruitment* (Apr. 12, 1999) <<http://www.uspto.gov/web/offices/ac/ahrpa/ohr/jobs/exam.htm>>. After 10 or 15 years, an examiner who has reached “Primary Examiner” status may earn \$72,000 to \$80,000. Telephone Interview with Jeff Kushan, Esq., Powell, Goldstein, Frazer & Murphy, Washington, D.C., and former Attorney-advisor with the Office of Legislation and International Affairs at the U.S. Patent and Trademark Office (Apr. 20, 1999); Cf. Jim Landers, *Perot-Backed Coalition Opposes Bill to Privatize U.S. Patent Office*, DALLAS MORNING NEWS, Sept. 8, 1998, at 1D (“[Then Commissioner of Patents Bruce] Lehman said privatizing the patent office would let him hire hundreds more patent examiners and pay them competitive salaries.”).

78. See Agnes Shanley, *You and Your Job: Shifting Career Gears Can Open New Doors*, CHEMICAL ENGINEERING 141, 141 (Dec., 1998) (reporting that average salary for entry-level chemical engineers was \$49,150 in 1998, versus an average of between \$95,700 and \$120,000 for management level engineers).

79. Telephone Interview with anonymous patent examiner (Feb. 1, 1999).

80. See Brenda Sandburg, *Patent Applications Flow Freely*, LEGAL TIMES, Feb. 22, 1999, at 12.

Economists have studied job performance when employees are assigned multiple tasks. Not surprisingly, if there are direct, "output-based" rewards for performing one task, but only diffuse, generalized rewards for performing the other(s), employees tend to devote most of their time to the directly-rewarded activity.⁸¹ Examples in the literature⁸² include salespeople who are also supposed to perform customer support. If their compensation is determined largely by sales commissions, they will tend to slight the customer support function. It requires large investments of resources to monitor and oversee their performance to prevent this effect. One suggestion of the literature is therefore that jobs should be separated by function where possible, so that there is less mixing of duties based on different compensation schemes. This thesis finds support in recent empirical work.⁸³

This logic applies readily to the job of structuring patent examiner incentives. There is a heavy burden on senior examiners. They are the primary training resource for new examiners. Yet they are subject to the same output-based compensation scheme as other examiners. This means they will tend to slight training. The obvious solution is to institute a thorough and effective training regime, under which senior examiners who provide training are directly compensated for the service. By all reports this has not been done. Much needs to be done to improve the quality of training that new examiners receive. If one assumes that the senior examiners are the most effective trainers, this simply adds to the reasons to scrap the existing output-based compensation system, or at least redesign it. One suggestion: routinely assign senior examiners to a training role, with a salary set at their average annual base salary-plus-bonus level for the past two years. (Obviously, they will have to wait at least two years

81. See generally Bengt Holmstrom & Paul Milgrom, *Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design*, 7 J.L. ECON. & ORG. 24 (1991). Holmstrom and Milgrom describe the importance of monitorability and employee incentives in jobs where employees (agents) are expected to perform multiple tasks. They present a model showing that separating tasks according to their monitorability characteristics allows the principal to give stronger incentives for tasks that are easy to measure, without fearing that the agent will substitute efforts away from harder-to-measure tasks. There are gains, in other words, from job designs that group hard-to-monitor tasks into individual jobs.

82. See, e.g., EDWARD P. LAZEAR, *PERSONNEL ECONOMICS* (1995).

83. See Trond Petersen, *Reward Systems and the Distribution of Wages*, 7 J. LAW, ECON. & ORG. 130 (1991); Holmstrom & Milgrom, *supra* note 81, at 24 (interpreting empirical studies).

between training stints.) This way they will not suffer economic loss from doing training.

B. Alternative Bonus Systems

The current bonus system is believed to skew incentives in favor of granting patents.⁸⁴ An obvious reform, then, is to change the bonus system. In general, the large literature on “personnel economics” ought to be brought to bear on the problem of designing a compensation system to advance the goal of a minimum acceptable error rate in patent issuances defined earlier. Here are some suggestions in this vein:

- Institute a tracking system to determine the “error rate” for examining groups and individual examiners, by assessing the percentage of patents issued by the group or examiner that are determined to be invalid in later court proceedings or reexaminations⁸⁵ on the basis of prior art that the examiner could have discovered; pay bonus compensation to groups and examiners whose error rates are lower than the office average or reach a pre-determined level acceptability;
- Outsource a selected sample of issued patents to a private-sector firm commissioned to determine the “error rate” on the date of issue; award bonuses to groups and examiners that beat the average error rate.

84. Consider these anonymous comments, posted to a patent examiners’ bulletin board:

You know what? I’m sick of finding ridiculous patents every time I look in my [files]. Part of the blame goes to the patent corps. We just don’t fight hard enough against the bull--- being shoveled by upper management. And of course, that is where the rest of the blame goes. It’s a system that’s burning up, and management just keeps adding fuel to the fire.

And why should you care? Hey, management pays you for good patents or bad, right? In fact, they pay you more for doing less. Why should you fight with management. Why reject?

Greg Aharonian, *A few patent examiners complain about patent quality*, INTERNET PATENT NEWS SERVICE (Jan. 28, 1999), available at <<http://lpf.ai.mit.edu/Patents/ipns/ipns-19990128.txt>>.

85. This would obviously necessitate a change in the current practice of giving a reexamination request to the same group or examiner that originally examined the application, but this is a good idea anyway given the normal human instinct not to admit a mistake.

C. Reforming Reexaminations: The Common Sense Case for Patent Oppositions

After a patent issues, anyone—including the patentee—can ask that it be reexamined.⁸⁶ Reexamination requests must be accompanied by a \$2,520 fee and a statement of the reason for the request.⁸⁷ By statute, the basis for reexamination is limited to certain types of prior art, in particular patents and printed publications.⁸⁸ And even if the request includes a new reference in one of these categories, reexamination will be initiated only if, in the opinion of the examiner, it raises “a substantial new question of patentability.”⁸⁹ Because reexamination is much cheaper than district court litigation⁹⁰—which can run anywhere from \$1 million to tens of millions of dollars for a patent case—it has obvious appeal. This explains the growth in reexamination requests reflected in Table 1.

Table 1. Annual Reexamination Filings⁹¹

| <i>Fiscal Year</i> | <i>Annual Filings</i> |
|--------------------|-----------------------|
| 1989 | 243 |
| 1990 | 297 |
| 1991 | 307 |
| 1992 | 392 |
| 1993 | 359 |
| 1994 | 379 |
| 1995 | 392 |
| 1996 | 418 |
| 1997 | 376 |

86. Reexamination proceedings are provided for by 35 U.S.C. §§ 301-07 (1998). Anyone, including the patentee, may ask the PTO to examine the patent in light of certain types of new prior art that was not considered during prosecution. If this raises a substantial new question regarding patentability, then the PTO grants a reexamination and determines whether or not the patent claims are still valid. *See MERGES, supra* note 7, at 1123-25.

87. *See* 35 U.S.C. § 302 (1998); 37 CFR § 1.20(c) (1998) (setting forth current fee). The Commissioner of Patents can also request a reexamination. *See MERGES, supra* note 7, at 1124.

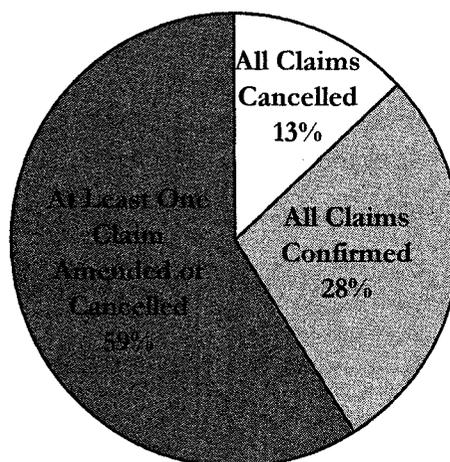
88. *See* 35 U.S.C. §§ 301-02 (1998). *See generally* MERGES, *supra* note 7, at 1124.

89. 35 U.S.C. § 303(a) (1998).

90. *See* H.R. REP. NO. 96-1307, at 3-4 (1980), *reprinted in* 1980 U.S.C.A.A.N. 6460, 6462-63.

91. *See* 1997 U.S. PAT. & TRADEMARK OFF. ANN. REP. 93 tbl.13; 1995 U.S. PAT. & TRADEMARK OFF. ANN. REP. 90 tbl.7; 1993 U.S. PAT. & TRADEMARK OFF. ANN. REP. 56 tbl.11.

What becomes of reexamination requests by third parties? The following chart, based on recent data,⁹² gives a summary:



Despite the growth in the number of reexamination requests, there is widespread dissatisfaction with the current system. This is especially true in comparison with European-style oppositions.⁹³ One commentator stated the case succinctly:

[T]he reexamination system implemented under this legislation has been underutilized and has not fulfilled its promise. In general, third parties have been unable to mount meaningful validity challenges under the reexamination system. For example, third parties have been limited in their ability to raise certain issues and adequately participate in the reexamination proceedings. In most instances, such parties choose to forego reexamination and instead await litigation in federal court. Consequently, while

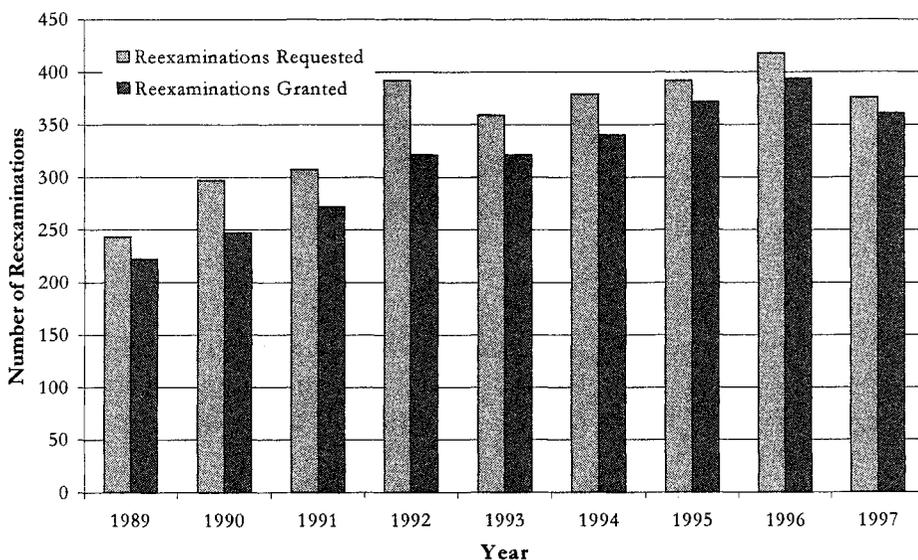
92. See Wayne O. Stacy, Note, *Reexamination Reality: How Courts Should Approach a Motion to Stay Litigation Pending the Outcome of Reexamination*, 66 GEO. WASH. L. REV. 172 (1997).

93. Oppositions, unlike reexaminations, are adversarial proceedings that allow for the introduction of physical evidence as well as the testimony of inventors and experts. In addition, oppositions occur early in a patent's life—they must be filed within nine months of issuance. This system finalizes patent validity earlier than the U.S. system, thereby benefiting patentees, potential infringers, and licensees. See MERGES, *supra* note 7, at 1131-34.

analogous systems in Europe and Japan have been effective in enhancing patent validity, the United States has struggled with an inadequate reexamination system.⁹⁴

Do the data bear this out? How does the U.S. reexamination system compare with Europe, which has a true opposition system?⁹⁵ Compare Table 2, which shows (1) the total number of opposition requests made to the European Patent Office, and (2) the percentage resulting in *total revocation* of the patent with Figure 3, which shows U.S. reexamination data.

Figure 3: U.S. Reexamination Data, 1989-1997⁹⁶



94. Soobert, *supra* note 76, at 66 (footnotes omitted).

95. The Japanese patent system also includes oppositions, but other differences between the U.S. and Japan, together with the general agreement over the efficiency of the European system, make Europe a better basis of comparison.

96. See 1997 U.S. PAT. & TRADEMARK OFF. ANN. REP. 93 tbl.13; 1993 U.S. PAT. & TRADEMARK OFF. ANN. REP. 56 tbl.11.

Table 2. European Opposition Data, 1994-1997⁹⁷

| Year | Oppositions Filed | Issued Patents Opposed (%) | Oppositions Resulting in Revocation (%) |
|------|-------------------|----------------------------|---|
| 1994 | 2,590 | 6.8% | not available |
| 1995 | 2,720 | 6.5% | 34.3% |
| 1996 | 2,600 | 6.2% | 33% |
| 1997 | 2,500 | 6.2% | 33% |

One can see immediately that the revocation rate is much higher in Europe compared to the United States—roughly 33%, versus 12% in the United States. Because it is difficult to quantify the effect of an opposition that does not result in a complete revocation, we can only speculate about the other 67% of oppositions.⁹⁸ It seems at least plausible, however, that the higher revocation rate implies something about the nature of the amendments in the cases where an opposition yielded a change in patent scope. To wit: we might well believe that oppositions lead to more substantial changes in patent scope than reexaminations do. The amendments made as a consequence of the high-quality information made available in an opposition would logically be more significant than in a reexamination, because (a) there are far more *categories* of prior art information available, and (b) the party collecting and presenting the information has a greater incentive to make it accurate and convincing.

Notice also the much higher *incidence* of oppositions in Europe, than reexaminations in the United States especially in light of the lower patent grant totals there.⁹⁹

97. See 1997 EUR. PAT. OFF. ANN. REP.15; 1996 EUR. PAT. OFF. ANN. REP.39; 1995 EUR. PAT. OFF. ANN. REP.44.

98. Cf. 1995 EUR. PAT. OFF., ANN. REP. 44 (Showing that, in approximately 65.7% of the cases, the patent was either maintained in amended form, or the opposition was rejected.)

99. The 2500 oppositions filed in Europe in 1995 were far in excess of the 376 U.S. reexamination requests. See 1995 EUR. PAT. OFF., ANN. REP. 15; 1997 U.S. PAT. & TRADEMARK OFF. ANN. REP. 93 TBL.13. In addition, there were far more opposition requests as a percentage of all patents issued the year before, or indeed of all outstanding patents in the system. By way of comparison, the European Patent Office granted 39,650 patents in 1997, versus 123,000 in the U.S. See 1997 EUR. PAT. OFF., ANN. REP. 15; 1997 U.S. PAT. & TRADEMARK OFF. ANN. REP. 84. Therefore, in Europe, 6.8% of issued

Creation of a coherent, efficient opposition procedure would be the ideal solution to a number of problems plaguing the current patent system. Short of this solution, recent proposals to reform reexamination in the U.S. are a step in the right direction.¹⁰⁰ They will, in the main, bring the U.S. practice more in line with Europe's. Though varied, recent proposals usually include some core components:

- More thorough participation of third-party requesters in the reexamination prosecution, e.g., presence at PTO interviews where crucial patentability advocacy takes place.¹⁰¹
- Possibility of appeals by third parties from adverse decisions by examiners during reexamination proceedings.¹⁰²
- Reform of the law regarding "staying" district court litigation during the course of a reexamination proceeding, a process which patentees can sometimes use for strategic delay.¹⁰³

Again, apart from the design details and implementation plan, the overall goal should be clear. We need to design a system that better taps

patents were opposed, whereas, in the United States, only 0.3% of patents were reexamined.

100. For a recent summary of current reform proposals—and a radical extension of them—see generally Soobert, *supra* note 76. Soobert's idea of creating a negative incentive to pursue oppositions, by in essence fining infringers who do not use oppositions, is interesting as a general idea but may not be workable in practice. It may result in over-monitoring of patents, in industries where firms might decide that taking the risks of later infringement lawsuits is worthwhile compared to spending current dollars on extensive monitoring of issued patents. A more positive incentive, such as the award of attorney fees to accused infringers who move to stay an infringement trial and later win an opposition, might be worth exploring.

101. Congress contemplated this practice in the 21st Century System Improvement Act:

While no statutory provision is added by this Act to address interviews conducted before the examiner during reexamination, it is intended that the Office, through rulemaking, will provide third-party requesters the right to participate in any examiner interview initiated by the patent owner or by the examiner, and that such interviews will be conducted under controlled conditions before the examiner and an additional, more senior, Office representative.

H.R. 400, 105th Cong. § 503 (1997).

102. See, e.g., S.507, 105th Cong. § 506 (1997).

103. See Stacy, *supra* note 92.

into patent validity information, much of which is in private hands. Until we get better information in the system, the quality of patents will not improve. Some may charge that oppositions will unduly favor big firms at the expense of independent inventors. Two points, however, must be kept in mind. First, the enhanced enforceability of patents that have survived oppositions is likely to be attractive to the investors who back small inventors; at any rate, these investors are likely to prefer the quicker and cheaper opposition system to expensive and protracted district court litigation. Second, companies that abuse the system by filing numerous and redundant requests for oppositions can be punished through such mechanisms as the award of attorney fees.

VI. CONCLUSION

In this paper, I have proposed some common-sense starting points to deal with the problem of business concept patents. In particular, I have tried to focus attention on determining an acceptable "error rate" for issued patents, with an eye toward reducing the number of invalid business concept patents that are actually issued. Second, I have refocused attention on the relationship between the PTO and private parties. The idea is to streamline the process so that it efficiently coordinates the efforts of both groups to achieve the socially desirable end: an appropriate expenditure for determining patent validity. The parties that suffer most if a company receives an invalid patent are that company's competitors. These parties also tend to have the best information about patent validity. Therefore, it is manifestly logical that they participate in the patent process as early and as thoroughly as possible.

Now, formerly "impossible" business concept and software patents are commonplace. The cost of the PTO's flawed granting and reexamination systems has become too high to ignore. We have fallen, like Alice, into a strange place where the normal rules do not apply, or have been inverted.¹⁰⁴ Lest vertigo get the better of us, forcing us to abandon all sense of logic and proportion, we must re-orient ourselves—take stock of the looking glass world of business concept patents, and see what we can do to restore some sense of order. The ideas in this article have been a step along this path.

104. *See supra* note 1.

OF GOVERNMENTS AND GOVERNANCE

By *A. Michael Froomkin*[†]

ABSTRACT

The Magaziner Report focuses on achieving short-term goals without giving sufficient consideration to long-term consequences affecting the structure of Internet governance and democracy in general. This overly pragmatic approach creates a paradoxical climate: overly-friendly to government intervention (in e-commerce regulation) while also overly-willing to defer to privatized governance structures (in other areas). As the recent World Intellectual Property Organization (“WIPO”) domain name/trademark process demonstrates, certain Internet governance processes raise several questions, not least discerning whether such processes include adequate notice and consultation. More traditional democratic processes, such as legislation and regulation, have routinized means of giving affected parties notice of pending decisions and of soliciting public comment. Other privatized governance processes may be equally or more legitimate, but not inevitably.

TABLE OF CONTENTS

| | | |
|------|---|-----|
| I. | INTRODUCTION | 618 |
| II. | WHOSE GOVERNANCE?..... | 619 |
| | A. Private Sector Autonomy?..... | 621 |
| | B. Local/National Government Rulemaking | 622 |
| | C. Harmonization | 623 |
| | D. Law “Reform” in Action..... | 625 |
| | E. Is E-Commerce Harmonization Necessary? | 629 |
| III. | CONCLUSION..... | 630 |

© 1999 A. Michael Froomkin.

† Professor, University of Miami School of Law and quondam Member, World Intellectual Property Organization Panel of Experts, Domain Name Process. Internet: froomkin@law.tm. Thanks for very helpful comments to Caroline Bradley, Joseph Froomkin, Patrick Gudridge, and Jonathan Weinberg. Except as otherwise noted, this paper seeks to reflect legal and technical developments as of March 6, 1999, the day it was delivered at the Berkeley conference on *The Legal and Policy Framework for Global Electronic Commerce: A Progress Report*. Readers are advised that the final version of the Interim WIPO Domain Name report discussed in this article should be available by late April 1999, at <http://wipo2.wipo.int/process/eng/processhome.html>.

I. INTRODUCTION

The second anniversary of *A Framework for Global Electronic Commerce*¹ ("Magaziner Report") provides an almost overdue occasion to reflect on U.S. government policy towards governance of an increasingly commercial Internet. In the two years since the Magaziner Report, a trickle of e-commerce has turned into a surge, and transactional floods loom. Few, however, foresaw the extent to which this flow of commercialization would create opportunities to erode, and perhaps ultimately overwhelm, the ad hoc governance structures that created and channel the Internet.

When viewed in light of these developments, the Magaziner Report seems so focused on achieving its short-term goals that it is insufficiently concerned with the long-term consequences of its recommendations for both the structure of Internet governance and of democratic government itself. The result of this overly pragmatic approach is to create a climate for too much governance—of the wrong sort.

Before turning to the implicit institutional theory found (or not found) in the Magaziner Report, I must disclose a source of potential bias. Having been appointed as the "public interest representative" to a Panel of Experts convened by the World Intellectual Property Organization ("WIPO") to advise it on conflicts between Internet domain names and trademarks, I am embroiled in a skirmish that forms part of the global war regarding the future of Internet governance. As an international body all too willing to take up the reigns of global governance, WIPO attempted to create global e-commerce friendly rules by a process that, left to itself, seemed likely to consist predominantly of meeting with commercial interest groups and giving little more than lip service to privacy and freedom of expression concerns. While the main theaters in the Internet governance struggle are clustered around the acronym soup of ICANN² and DNSOs,³ I have been

1. WILLIAM J. CLINTON & ALBERT GORE, JR., *A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE* (1997), available at <<http://www.ecommerce.gov/framework.htm>> (discussing the need for a set of globally recognized commercial law rules) [hereinafter FRAMEWORK].

2. ICANN is the Internet Corporation for Assigned Names and Numbers. ICANN owes much of its authority to an agreement with the U.S. Department of Commerce. See *Memorandum of Understanding Between the U.S. Department of Commerce and ICANN* (visited Mar. 3, 1999) <<http://www.ntia.doc.gov/ntiahome/domainname/icann-memorandum.htm>>.

3. DNSO is an acronym for "Domain Name Supporting Organization." The Commerce Department's White Paper on Management of Internet Names and Numbers—a Statement of Policy dated June 10, 1998—suggested that ICANN "could rely on separate,

resisting what one might hyperbolically call an attempted “trademark grab.”⁴ The experience risks turning me into a nationalist. If nothing else, it has reinforced an already strong belief that governance structures matter at least as much as the content of any ephemeral set of rules. As Jean Monnet put it, “nothing is possible without men [sic], nothing is lasting without institutions.”⁵ Attention to institutional issues, however, is where the Magaziner Report now appears most lacking.

II. WHOSE GOVERNANCE?

It may seem odd to accuse the Magaziner Report of insufficient attention to institutional issues when the document begins with a declaration of principles which on first reading appear to suggest close attention to the institutional design of e-commerce regulation. These five principles are:

- The private sector should lead.
- Governments should avoid undue restrictions on electronic commerce.
- Where governmental involvement is needed, its aim should be to support and enforce a predictable, minimalist, consistent and simple legal environment for commerce.
- Governments should recognize the unique qualities of the Internet.
- Electronic commerce over the Internet should be facilitated on a global basis.⁶

diverse and robust name and number councils responsible for developing, reviewing, and recommending for the board’s approval policy related to matters within each council’s competence. Such councils, if developed, should also abide by rules and decision-making processes that are sound, transparent, protect against capture by a self-interested party and provide an open process for the presentation of petitions for consideration.” Management of Internet Names and Addresses, 63 Fed. Reg. 31,741, 31,750 (1998), available at <http://www.ntia.doc.gov/ntiahome/domainname/6_5_98dns.htm> [hereinafter *White Paper*].

4. For an explanation, see A. Michael Froomkin, *A Critique of RFC 3* (last modified Feb. 21, 1999) <<http://www.law.miami.edu/~amf>>. See also Pamela Samuelson, *The Copyright Grab*, WIRED, Jan. 1996, at 134 (for perspective on how intellectual property rights holders seek to manipulate legal rules to their pecuniary advantage).

5. JEAN MONNET, MEMOIRES 360 (1976) (“Rien n’est possible sans les hommes, rein n’est durable sans les institutions.”).

6. FRAMEWORK, *supra* note 1, *Principles* 2-3.

Despite these declarations of principle, in hindsight it seems clear that what is missing from the *Framework for Global Electronic Commerce* is ... a framework. There is a vision in the Magaziner Report that might merit the word “strategic”—there is some sense that we are *here* today and we wish to navigate to be *there* tomorrow—so it could be unfair to say that the Magaziner Report is consumed by tactics. It is fair to say, however, that the Magaziner Report is consumed by short-term policies and fails to grasp the consequences of the means proposed to achieve its short-term ends for long-term global governance.

The tensions implicit in the Magaziner Report’s approach to governance become evident when one considers the role contemplated for the private sector. The promise of no “undue” restrictions by government, but rather “support” and “minimalist” rules, just sufficient to “facilitate” e-commerce “on a global basis” sounds uncontroversial.⁷ On the one hand, the private sector is portrayed in its heroic mode, needing only to have moribund rules removed to allow its unleashed animal spirits to carry the day. In effect, the private sector rules, or should rule. Yet, on the other hand, the private sector needs to be supported, cosseted, and to have rules optimized for it on a global basis. Thus governments have a three-faceted role. First, they duly promulgate restrictions on e-commerce when needed. Second, they facilitate e-commerce by creating simple, predictable rules. Third, they join together to do more of the same. It might be possible to keep both the promise of minimalist rules and the promise of activist intervention to make the world safe for e-commerce, but given the tenor of the Magaziner Report’s general recommendations, it seems most likely that the interventionist tendency will win in any conflict between the two.

Indeed, while the rhetoric of the Magaziner Report exalts the private sector and emphasizes its autonomy, the action program in the Magaziner Report reveals a different view. Despite the free-market tone of the first principle, the Magaziner Report is far from a libertarian document. Instead, the Magaziner Report opens up the possibility of a host of new rules emanating from a variety of sources. The Magaziner Report’s strategic goal is to further the adoption of what its authors believe to be the right rule set—one that creates an optimal climate for e-business—and the authors are not at all doctrinaire about where those rules should come from. Thus, the Magaziner Report proposes to work with and through whatever institutions look most poised to advance the adoption of the right rules. Sometimes this means private sector autonomy, other times it means

7. *Id.*

local/national rule-making (akin to “subsidiarity”⁸), but most often it turns out to mean some form of globalized rule-making. This is considerably more government action than one might expect from a regime of “minimalist” rules.

A. Private Sector Autonomy?

Despite the general tendency towards globalized rulemaking, there are three areas where the Magaziner Report makes a fairly strong stand for private sector autonomy. First, it gives strong support to private efforts to address basic Internet governance issues such as the problems of allocation of domain names, although in practice it turns out that these efforts are not all that private.⁹ Second, the Magaziner Report supports the continued development of voluntary technical standards, which in Internet terms is akin to supporting motherhood. And, third, the Magaziner Report endorses a self-regulation regime for privacy principles in general defiance of the stronger measures suggested by the European Privacy Directive—although it warns that if self-regulation is not forthcoming, the government may find itself politically obligated to regulate. Here, private sector autonomy takes precedence over privacy.

In contrast to this vision of private sector autonomy, consider the role envisaged for the private sector in electronic payments. The Magaziner Report admits that in the short term the private sector must lead in electronic payment systems because the technology is changing too quickly for the government to regulate effectively. “In the near term,” therefore, the government role is limited to “case-by-case monitoring of electronic payment experiments.”¹⁰ This is only temporary, however, for, “[f]rom a longer term perspective, ... the marketplace and industry self-regulation alone may not fully address all issues. For example, government action may be necessary to ensure the safety and soundness of electronic payment systems, to protect consumers, or to respond to important law enforcement objectives.”¹¹ Of course, the private sector will be consulted to “ensure that governmental activities flexibly accommodate the needs of the emerging marketplace.”¹² Consider also the role foreseen for con-

8. Subsidiarity is the devolution of responsibility to smaller political units in the context of a federal system.

9. Compare FRAMEWORK, *supra* note 1, § 4, at 11 (discussing domain name issues), with *White Paper*, *supra* note 3 (calling on a public body, WIPO, to make recommendations on domain name management).

10. FRAMEWORK, *supra* note 1, § 1, at 6.

11. *Id.*

12. *Id.*

tracting parties on the Internet. The Magaziner Report envisions a world where freedom of contract is king, and parties will be masters of their contract, right down to choice of law and forum. Nevertheless, parties should expect to bargain in the shadow of a single, world-wide, agreed U.C.C.-like set of rules to be developed at the international level.

Ultimately, in the Magaziner Report vision, the private sector will not lead; it will instead hold sway within the confines defined for it, its traditional role in a mixed capitalist economy. Strangely, it may turn out that in the Magaziner Report vision of the near future, certain private parties will find their greatest empowerment and autonomy not in the marketplace, but in the bargaining process by which the new global rules shaping that marketplace will formed.

B. Local/National Government Rulemaking

The Magaziner Report seems to find almost no role for state/local governments in the regulation of e-commerce. They do not appear often in the document, except when they are to be discouraged from levying Internet taxes. One might say that this absence follows from the international nature of the Internet. The Magaziner Report implicitly argues that if e-commerce rules are to be consistent, then policies need to be made at least at the national level, and perhaps at the global level.¹³ Given that in the U.S. the primary regulatory authority for the law of sales has tended to rest with the states (with harmonization enhanced by coordinated law projects such as the U.C.C.), this would represent a larger shift towards federal or international rulemaking than perhaps the Magaziner Report lets on.

In contrast, national rulemaking figures in the Magaziner Report in two ways. First, there are a few areas which are identified as suited for straightforward national legislation. Examples include legislation on server/provider liability,¹⁴ fraud prevention in general,¹⁵ and the regulation of cryptography.¹⁶ We must accept cryptography control as *sui generis*, though headed for the footnotes of history.¹⁷ But how exactly Internet fraud became primarily a matter of federal, and even international, jurisdiction is more asserted than explained. Securities regulation excepted, most law relating to commercial fraud, like most contract law, has traditionally been a state responsibility. Perhaps the Magaziner Report's tilt

13. *See id.* § 3.

14. *See id.* § 4, at 9.

15. *See id.* § 8, at 19-20.

16. *See id.* § 6, at 15.

17. *See infra* text accompanying notes 33-35.

towards federal/international jurisdiction reflects the current U.S. reality that the majority of law enforcement and prosecutors with experience in Internet-related crime work for the larger, usually federal, departments. Some matters, such as the issuance of rules to combat online securities fraud, are properly national (although there is some concurrent regulatory jurisdiction at the state level).

Garden-variety fraud that moves off-shore presents jurisdictional problems, and new opportunities for international cooperation among law enforcement bodies, but it is less evident that it requires new substantive rules. Nevertheless, an argument might be made to justify this tilt: arguably because the Internet makes every consumer transaction feel equally local, and because absent a robust digital signature infrastructure consumers are not able to verify the nationality of a merchant,¹⁸ potentially fraudulent consumer transactions are now more similar to potentially fraudulent securities transactions. Consumer protection law therefore must make provisions at the national and international levels for cross-border fraud, just as the securities regulation regime has done. However, no such argument is found in the Magaziner Report, and it is vulnerable to the counter-argument that the best cure is the provision of technical means such as a robust digital signature infrastructure¹⁹ (sometimes called a "public key infrastructure" or PKI) rather than national or supra-national rulemaking.

It is easier to understand how a concern with the preservation of values of free expression leads the Magaziner Report to the assertion of national primacy when it comes to content controls and the regulation of both political and commercial speech. Indeed, the Magaziner Report states that not only will the U.S. retain full autonomy on matters of free expression, but that the U.S. will try to spread the gospel of the First Amendment throughout the world.²⁰ This deserves great praise because freedom of expression is far too important to be risked to the doubtful processes proposed for electronic commerce rulemaking.

C. Harmonization

Second, and perhaps even more important, nations figure as both participants and subjects in the harmonization of international law. Through-

18. See generally A. Michael Froomkin, *The Essential Role of Trusted Third Parties in Electronic Commerce*, 75 OR. L. REV. 49 (1996), available at <<http://www.law.miami.edu/~froomkin/articles/trusted.htm>>.

19. See *id.* (explaining function and uses of a public key infrastructure).

20. See FRAMEWORK, *supra* note 1, § 8, at 18.

out the Magaziner Report one finds calls for international cooperation of various sorts, for action by international organizations, and for legal standardization generally.

International legal harmonization happens in a variety of ways, ranging from highly decentralized to top-down rulemaking. A few examples illustrate the spectrum of means that might be available.

- The most decentralized form of harmonization occurs when norms, usages of trade, *lex mercatoria*, or the like spontaneously develops within a (usually specialized) transnational (usually commercial) community.
- The second most decentralized form of harmonization happens when one jurisdiction's law becomes the de facto rule for another place, perhaps due to regulatory arbitrage. For example, if a country with excellent Internet connectivity has a policy of allowing full freedom of expression and anonymous Internet access, this policy will have effects on every other nation that chooses to allow a full Internet feed.²¹
- A third form of harmonization occurs in the context of regulatory competition, when one jurisdiction chooses to copy another's rules, whether as part of a race to the bottom, or a struggle to the top.
- Governments participate in a fourth type of harmonization when they engage in communal law reform projects, e.g. under the auspices of the United Nations Commission on International Trade Law ("UNCITRAL"), that produce model laws that are then presented to states for their enactment, much as the Commissioners on Uniform Laws in the United States produce model legislation that is presented to the states for their approval.
- Fifth, supra-national bodies are often designed and empowered to harmonize the law of member states, with the European Union and the North American Free Trade Agreement ("NAFTA") being major examples.
- And finally, although sometimes cumbersome to enact, international and especially multilateral treaties potentially are a very powerful source of legal harmonization.

21. See generally A. Michael Froomkin, *The Internet As a Source of Regulatory Arbitrage*, in BORDERS IN CYBERSPACE: INFORMATION POLICY AND THE GLOBAL INFORMATION INFRASTRUCTURE 129 (Brian Kahin & Charles Nesson eds., 1997), available at <<http://www.law.miami.edu/~froomkin/articles/arbitr.htm>>.

There are innumerable differences between each of these modes of harmonization, and of course not every one will be available, much less appropriate, to serve as the solution to any given perceived problem. The key differences, however, are the extent to which these processes are democratic, and the extent to which they are subject to capture. Although a great deal depends on the nature of the issue and the circumstances, on balance it seems that among the options subject to governmental control, the option of presenting model laws to a legislature for adoption is best calculated to produce legal harmonization without sacrificing basic democratic values.

Let me quickly admit that we do not begin from the highest baseline, and that motives are sometimes purer than the word "capture" may suggest. No process of human decision making, and certainly none that involves institutions, can ever be perfectly democratic, or completely immune from capture. And certainly the lawmaking process in the United States shows signs of suffering in this department. Indeed, at a conference organized at the University of California, Berkeley less than a year ago, many joined in an effort to explore the ways in which proposed Article 2B of the UCC would favor certain commercial interests over other participants in the market.²² That experience also serves to remind us that one person's "capture" is another person's sincere belief that optimalities are found in different places. But that experience also serves to remind us of the salutary effect of bringing one's proposal before a jury of one's professional peers—and then having to try to get it through a real legislature.

D. Law "Reform" in Action

Even in a world of shades of gray there can be better and worse. International standard-making, although not beyond capture, on the whole fits into the "better" category. The Internet Engineering Task Force ("IETF") process of long-winded discussion and peer review retains virtues even as the corporate vice-presidents are crowding into the process. In contrast, most of the processes of international harmonization proposed in the Magaziner Report strike me, on the whole, as likely to produce two kinds of processes that both fall into the "worse" category: (1) the traditional

22. See Symposium, *Intellectual Property and Contract Law in the Information Age: The Impact of Article 2B of the Uniform Commercial Code on the Future of Transactions in Information and Electronic Commerce*, 13 BERKELEY TECH. L. J. 809 (1998); Symposium, *Intellectual Property and Contract Law for the Information Age: The Impact of Article 2B of the Uniform Commercial Code on the Future of Information and Commerce*, 87 CALIF. L. REV. 1 (1999).

multi-lateral treaty process, and (2) an international version of the process that produced Article 2B, without legitimating ratification by legislatures.

For all their virtues, treaties are not democracy at its finest. At best, treaties are negotiated by unelected delegates of elected officials; they get to negotiate with their counterparts and with the unelected delegates of despots. The ability of elected officials outside the executive branch, and all but the best-informed and well-financed interest groups, to influence the negotiation process is ordinarily attenuated—although in rare cases far from zero as the recent career of Senate Foreign Relations Chairman Jesse Helms demonstrates.²³ Treaties come to legislatures in a form which largely is not amendable.²⁴ In the U.S., treaties are subject to ratification in only one House; at times they are a means for canny administrations to get Congressional agreement to things that would never have passed both houses if seen to be of U.S. origin. Some will call the practice of using international processes to achieve results unachievable by ordinary legislation, such as the Digital Millennium Copyright Act, a form of high statecraft. Even if true, when it reaches the regulation of the ordinary commercial life of the nation, it is statecraft with significant costs: treaties are costly to break and hard to amend.

Purely private lawmaking is near the other extreme of the continuum of harmonization mechanisms. An international analog of the private lawmaking process that produced Article 2B may be even more discomfiting: when the topic is international harmonization the meetings are often farther away, the air tickets and hotels cost more, and even more people are forced to disenfranchise themselves as the process drags on and their time and money become exhausted. Only those with fee-paying clients, or with obsessions, can stay the course, and yet it is only a matter of time before the twin cries of laches and estoppel are heard in the land. The result gains democratic legitimacy when adopted by a legislature, but a flawed process is nonetheless troubling because the proposals may come to the legislature with a patina of legitimacy they may not deserve.

The WIPO domain name system/trademark (“DNS/TM”) process in which I have been an “Expert” participant is worse still, because the results will never have to be presented to a legislature. WIPO is a body formally composed of its member states, but blessed with an energetic, intel-

23. See, e.g., Richard L. Berke & Steven Lee Myers, *In Washington, Few Trifle With Jesse Helms*, N. Y. TIMES, Aug. 2, 1997, § 1, at 1 (noting and describing power wielded by current Chairman of Senate Foreign Relations Committee).

24. Nations can ratify a treaty with reservations attached, but there are diplomatic costs to this practice.

ligent, and surprisingly well-financed Secretariat. In part at the request of the U.S. government,²⁵ WIPO took on the task of crafting recommendations relating to intellectual property concerns caused by the increasing use of domain names as a marketing tool. The recommendations will be passed on to the fledgling Internet Corporation for Assigned Names and Numbers ("ICANN"), a private not-for-profit California corporation charged by the U.S. government with taking over key coordination functions for the technical management of the Internet. ICANN is not formally obligated to accept WIPO's recommendations, but there is likely to be considerable political pressure for it to take some or all of them, and ICANN currently is not formally accountable to anyone.

There is no denying that the DNS/TM problem is complex. It involves conflicts caused by trying to map a trademark system that is both geographic and sectoral onto a domain name space that is world-wide and has only one .com. In addition, there are conflicts between trademark owners and those with other legitimate interests in a domain name—interests that range from nicknames and surnames to criticism and parody. In addition, there are conflicts caused by speculative behavior and hoarding (sometimes termed "cybersquatting" or "cyberpiracy"), or outright attempts to deceive by passing off one site as associated with another's brand.

WIPO responded to this challenge by proposing that ICANN use its (arguable) leverage over the bodies that will control the databases of mappings between Internet domain names and IP numbers to impose a series of contractual duties on all users of those databases.²⁶ Details aside, the key point is that WIPO proposed a series of contracts of adhesion that would result in every registrant in .com, .net, or .org having to agree to a loser-pays arbitration under substantive rules that likely would differ from the laws applied by a competent court. No treaty or legislation would be required. As the arbitrators would be instructed to supplement applicable national law with certain "principles" identified by WIPO, some results likely would differ from what a court would do. At this writing, the final report has not been written, and there remains at least some hope that it will be much changed from an Interim Report that I believe is biased in favor of intellectual property rights holders.²⁷

25. See *White Paper*, *supra* note 3.

26. See World Intellectual Property Organization, *The Management of Internet Names and Addresses: Intellectual Property Issues* (Dec. 23, 1998) <http://wipo2.wipo.int/process/eng/rfc_html> (Interim Report of the WIPO Internet Domain Name Process) [hereinafter *WIPO Interim Report*].

27. See Froomkin, *A Critique of RFC 3*, *supra* note 4.

The WIPO process represents an innovative experiment in negotiation between a United Nations body and the private sector. Once having authorized the initial process, the member states have been conspicuous by their absence, at least as a formal matter, except as commentators in response to the various “requests for comments” authored by the Secretariat staff. Regardless of the merits of the WIPO proposals, there can be little debate that the public participation in the process has been dominated by intellectual property rights holders and their lawyers and trade associations.²⁸ Similarly, the Secretariat staff appear to be very sincerely committed to WIPO’s mission of the promotion of intellectual property rights—so much so that to even think about “capture” almost seems besides the point.

The WIPO DNS/TM process has certainly been public in a formal sense, with a series of meetings around the world, and web pages displaying documents and public comments. But public participation has been low for a number of reasons, including poor publicity outside the intellectual property community, and especially the competition for the attention of the relatively small number of people focused on the issue of Internet governance. Most of them understandably have focused on decisions relating to the structure of ICANN rather than on a merely advisory report, even one likely to be influential. Turnout at the public hearings I have attended has been small—usually under 100 and sometimes about 50, and (with the exception of the Washington D.C. event that followed a publicity campaign I organized) comprised almost entirely of trademark lawyers or Internet service providers.²⁹ There have also been over 150 e-mailed comments.

The dearth of consumer representatives, public interest groups, and citizens groups participating in the WIPO process should serve to remind us all of the many reasons why we entrust major aspects of social policy making to elected officials. Legislatures and national executives are not paragons of rectitude. But when they are elected, the same political proc-

28. Since I delivered this paper in Berkeley the balance has been somewhat redressed, in substantial part due to the written comments to WIPO by several members of the conference audience. See, e.g., Kurt Opsahl, *Law Professors, Academics, Students, Attorneys and Industry* (last modified Mar. 19, 1999) <http://wipo2.wipo.int/dns_comments/rfc3/0164.html> (submitting letter signed by more than 60 opponents of RFC 3).

29. I base this on my attendance at the Toronto, Rio de Janeiro, Brussels and Washington, D.C. second round consultations. Transcripts of these hearings, and the Singapore and Dakar hearings also, will be available at <<http://wipo2.wipo.int/process/eng/consult2.html>>.

ess that may make them over-solicitous to those bearing campaign contributions imposes some form of accountability to the public at large. Equally important, presentation of a matter to an elected official is a way of putting a question onto the public agenda. WIPO has a very nice set of web pages that lay out the issues at stake in the DNS/TM process and lay out the schedule for its public meetings in various world capitals.³⁰ There is no particular reason, however, to assume that anyone is necessarily going to know that those web pages are there, or would necessarily visit them. In contrast, a hearing in front of a subcommittee, a vote by a house of Congress, or even a publication of a proposed rule in the federal register by an unelected bureaucrat, would serve to put the public (in one country) on notice in a tolerably effective way—at least in a routine and knowable way—of the rules that someone proposes to lay down upon them.

Indeed, it is not obvious that all the relevant portions of governments understood what was going on. WIPO sent notices of its proposals to every one of its member states, but one suspects from the responses received that these were directed at the patent and trademark offices with which WIPO ordinarily corresponds.³¹ Whether these notices were then circulated to other departments is hard to ascertain.³²

E. Is E-Commerce Harmonization Necessary?

The world envisioned by the Magaziner Report, at least in the parts that use the rhetoric of the leading role of the private sector, is one in which the contractual autonomy of parties extends to making enforceable contracts with choice of law and choice of forum. Competent parties with the ability to pick law and forum may not need as much international legal harmonization. Whether harmonization is necessary becomes doubly important if the processes used for achieving harmonization create or deepen a democratic deficit.

Undoubtedly, there are areas where some harmonization is required. Without basic plumbing such as the recognition and enforcement of electronic contracts, online commerce cannot flow safely. After the basics,

30. See, e.g., *World Intellectual Property Organization Internet Domain Name Process* (visited Apr. 6, 1999) <<http://wipo2.wipo.int/process/eng/processhome.html>>.

31. Comments and responses are available at <http://wipo2.wipo.int/dns_comments/rfc3/index.html>.

32. At least within the U.S. government, consultation was imperfect. Eric Menge, Office of Advocacy, U.S. Small Business Administration, Statement at the WIPO consultative meeting (Washington, D.C.) (Mar. 10, 1999) (stating that he had only learned of the proposals a week earlier).

however, clarity will ordinarily suffice if parties have the autonomy to contract around legal impediments. From a government's point of view, the major areas that require additional harmonization are those where risks exist that citizens will engage in undesirable forms of regulatory arbitrage or that consumers will be preyed upon by the unscrupulous. Governments and all those who consider taxes the price of civilization have a common interest in controlling opportunities for tax avoidance and evasion. Similarly, governments concerned about various law enforcement issues, especially the control of money laundering, may have common cause to make inter-jurisdictional agreements for the regulation of electronic cash and transborder deposit-taking. And, considerations of consumer protection (which the private sector tends to lean against rather than lead) may argue for rules requiring transparency or rules regularizing expectations.

Other forms of regulatory arbitrage suit U.S. interests. The Magaziner Report correctly identifies the enhancement of freedom of expression as something too valuable to risk to a multilateral process designed to create some international set of content controls. Given our federal system, however, it seems a little odd to find not only that subsidiarity is mostly absent from the Magaziner Report, but that the contribution of nations as "big labs of democracy" is not recognized.

III. CONCLUSION

Hindsight, of course, is a beautiful thing. Two years ago, when the Magaziner Report first issued, I saw none of this. Instead I was excoriating the Magaziner Report for failing to take a principled stand on reform of cryptography regulation. At the time, I thought that one of the best things the government could do for e-commerce would be to reverse its twin policies of manipulating markets to favor cryptographic "key escrow" and its long-standing ban on the export of meaningful cryptography. Unable to make headway against the Administration consensus in favor of those policies,³³ the Magaziner Report echoed the Janus-like pronouncements that the Administration supported strong cryptographic security measures, so long as one did not have enough to ensure true security and went on to

33. See generally A. Michael Froomkin, *The Metaphor is the Key: Cryptography, the Clipper Chip and the Constitution*, 143 U. PENN. L. REV. 709 (1995), available at <<http://www.law.miami.edu/~froomkin/articles/clipper.htm>>; A. Michael Froomkin, *It Came From Planet Clipper*, 1996 U. CHI. L. FORUM 15, available at <http://www.law.miami.edu/~froomkin/articles/planet_clipper.htm>.

other things.³⁴ Today, as strong consumer cryptography spreads around the globe, the Magaziner Report's timidity on this issue almost seems the better part of valor. With even France proposing to liberalize its rules to allow unrestricted use of 128-bit cryptography,³⁵ it seems increasingly likely that U.S. policies restricting the spread of strong consumer cryptography are becoming irrelevant. Strong cryptography may not be built into Windows 2001, but it will be in the operating system of the near future.

The same social process that is making U.S. cryptography policy increasingly irrelevant internationally suggests that top-down regulations of the type proposed in the Magaziner Report may become irrelevant too. If that scenario is at all real, it may be one more reason to balk at processes that do not have a democratic pedigree. If the private sector is poised to lead from the bottom up, perhaps we should celebrate that, not stomp on it, even (especially) if the redistributive effects of this evolution disadvantage large, established parts of the corporate sector at the expense of small companies arising in the new web-based economy. But it is one thing to celebrate market-driven outcomes (corrected for market failures), and to value market-making technical standardization. It is quite another thing to tolerate private sector leadership when it clothes itself in the guise of "bottom-up rulemaking" but actually seeks to use government or government-like power to lock in advantages enjoyed by established firms, often at the expense of consumers or new competitors.

The real challenge comes in telling legitimate processes apart from the others. True "bottom-up" outcomes are entitled to respect. Not every privatized rulemaking procedure carries the same legitimacy, however. The WIPO process in which I participated involved limited consultation, and opaque decision-making as the reports and recommendations were all written by the Secretariat staff in secret. It featured what appeared from my admittedly partisan perspective to be disproportionate input by trademark holders (as opposed to actual and would-be domain name holders without trademarks), especially in the period before I began to kick up a

34. True security would make it difficult or impossible for law enforcement and intelligence agencies to intercept communications when they, and a court, believe it necessary.

35. See *Conférence de presse de Monsieur Lionel JOSPIN, Premier ministre, à l'issue du Comité interministériel pour la société de l'information Hôtel de Matignon* (Jan. 19, 1999) <<http://www.premier-ministre.gouv.fr/PM/D190199.HTM>>; see also *Décret n° 99-199 du 17 mars 1999 définissant les catégories de moyens et de prestations de cryptologie pour lesquelles la procédure de déclaration préalable est substituée à celle d'autorisation* (Mar. 19, 1999) <<http://jya.com/decret031799.htm>>.

fuss. Rather than be presented to one or more legislatures for ratification, the result will be presented to ICANN, a private non-profit corporation, and if accepted by ICANN will be imposed contractually on all registrants in global top-level domains. This contrasts with the adoption of a technical standard which, at least until the network effects kick in, attracts adherence without the need for any form of external pressure or coercion.

The WIPO process also contrasts with democratic processes. Democratic societies, and especially the U.S., have evolved elaborate techniques for giving notice of upcoming decisions, and making it possible to spot capture in the lawmaking process. Various countervailing advocacy and special interest groups monitor the legislative process seeking to represent the interests of the public or segments of it. Relative to the size of the need, the equivalent public sector is much more attenuated at the international level, and monitoring is in any event much more difficult as institutions are less transparent and much more diverse and spread over the world. As a result, even active citizens and legislators are less able to weigh the results of internationalized rulemaking processes, not least because they cannot have confidence that if the process was seriously deficient someone would blow the whistle. Even the existence of a self-proclaimed whistle-blower is not enough: there is no particular reason to believe that the people claiming a substantive or procedural flaw have a way of getting the attention of relevant auditors, nor for thinking that the auditors have any way of judging the credibility of the source even if they happen to hear the communication.

The difference between the governance procedures endorsed by the Magaziner Report, of which the WIPO process is but one example, and true bottom-up rulemaking can be analogized to the difference between a proprietary standard and Open Source software.³⁶ Like a manufacturer choosing a proprietary standard, in the WIPO process a single body makes decisions, in relative privacy, after soliciting the degree of input it feels is appropriate. The outcome of the process ought then to be launched on the marketplace of ideas. Instead, the proprietary legal standard threatens to become dominant in its category because ICANN will present consumers with a *fait accompli*. In its Interim Report, WIPO proposes that ICANN impose WIPO's ideas on registries, who would in turn impose them on registrars who would in turn be required to impose them on domain name

36. On Open Source software, see *GNU General Public License* (visited Apr. 6, 1999) <<http://www.gnu.org/copyleft/gpl.html>> (text of model license); Ira V. Heffan, *Copyleft: Licensing Collaborative Works in The Digital Age*, 49 STAN. L. REV. 1487 (1997). I owe the analogy to a conversation with Patrick Gudridge.

registrants.³⁷ This is more lock-in than is allowed in product markets; even Microsoft's efforts to lock in a large market share by making Windows the required default option on all personal computers sold by a given manufacturer could be reversed by the consumer. It also contrasts with the propagation of open source software, which can freely be copied or modified by anyone so long as the source is open and it is licensed for free. The terms of the open source license are imposed on subsequent designers who choose to incorporate or user of the code, but the decision whether to use the code is left to them.

Undoubtedly it was not the intention of the Magaziner Report to endorse proprietary private lawmaking. We accept proprietary legal solutions when they emanate from a legislature, whatever the source of the original draft. Legal solutions pioneered by judges we often dub "common law"³⁸ and count on competition between states to produce the best, or a range of acceptable solutions. Other models still must prove themselves.

37. See *WIPO Interim Report*, *supra* note 26, ¶¶ 57, 101, 142, 221.

38. Matters are more complex in the statutory context, where we sometimes dub judicial creativity "common law-like"—and yet more complex in the constitutional context.

PROGRESSING TOWARDS A UNIFORM COMMERCIAL CODE FOR ELECTRONIC COMMERCE OR RACING TOWARDS NONUNIFORMITY?

By *Maureen A. O'Rourke*[†]

ABSTRACT

The Magaziner Report encourages the development of a consistent commercial law environment against which electronic commerce transactions may take place. The author considers the current legal landscape, noting that while many efforts are underway to codify aspects of electronic commerce, these efforts are piecemeal in nature and may lead to the very lack of uniformity against which the Magaziner Report counsels. The author then briefly considers what lessons may be learned from the drafting history of the original U.C.C. as well as proposed Article 2B (now the Uniform Computer Information Transactions Act) governing transactions in computer information. She argues that Article 2B could benefit from some synthesis with efforts by the European Union and that the Clinton Administration should focus on identifying and harmonizing mandatory rules of law as it moves from broad statements of principle to practical implementation of a legal regime for electronic commerce.

TABLE OF CONTENTS

| | | |
|------|--|-----|
| I. | INTRODUCTION | 636 |
| II. | CODIFICATION EFFORTS..... | 638 |
| III. | OBSTACLES TO UNIFORMITY | 641 |
| IV. | THE LEGACY OF LLEWELLYN: LESSONS FOR THE DRAFTERS OF ARTICLE 2B AND OTHER LAWMAKERS..... | 645 |
| | A. A Brief History of Article 2..... | 645 |
| | B. Article 2B and Information Transactions..... | 648 |
| | C. Article 2B and Intellectual Property Law | 649 |
| | D. Article 2B, the "Agreement," and Consumer Protection | 651 |
| | E. Summary: Striking a Balance Between Uniformity and Flexibility | 653 |
| V. | CONCLUSION..... | 657 |

© 1999 Maureen A. O'Rourke.

† Professor of Law, Boston University School of Law. Thanks to James Molloy and to Professors Pamela Samuelson and Mark Lemley for inviting me to participate in the Berkeley Conference on the Legal and Policy Framework for Global Electronic Commerce and to write this paper as part of that participation.

I. INTRODUCTION

As the Clinton Administration recognized in its 1997 *Framework for Global Electronic Commerce* ("Magaziner Report" or "Report"), the Internet is transforming all aspects of life—from changing the way we think about what constitutes a community, to how we conduct research, to how we enter into and perform commercial transactions.¹ As its official title suggests, the Magaziner Report as well as the Presidential Directive accompanying it both emphasize the latter transformation. They consider primarily how the Internet affects global trade in goods, services and information and what the government's role should be in facilitating that trade.

The Report's premise is fairly straightforward. The Internet is characterized generally by both low barriers to entry and low transaction costs. These and other qualities, like the Internet's dynamic technology and global nature, should lead to a highly competitive market resulting in greater choice for consumers at lower prices than in conventional markets. The animating philosophy behind the Report is that this is a desirable state and therefore the law should be drawn to encourage parties to engage in electronic commerce. In the Report's view, the law may do this best by establishing technology-neutral, framework principles to provide a certain, uniform legal environment against which the market may work.² Establishing such a framework would include both dismantling existing legal barriers to electronic commerce and enacting as necessary new rules to support the use of technology. Moreover, consistent with the historical

1. WILLIAM J. CLINTON & ALBERT GORE, JR., A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE (1997), available at <<http://www.iitf.nist.gov/elecomm/ecommm.htm>> [hereinafter FRAMEWORK]. The Magaziner Report comments:

No single force embodies our electronic transformation more than the evolving medium known as the Internet.... [T]he Internet has emerged as an appliance of every day life.... Students across the world are discovering vast treasure troves of data via the World Wide Web.... Citizens of many nations are finding additional outlets for personal and political expression. The Internet is being used to reinvent government and reshape our lives and our communities in the process. As the Internet empowers citizens and democratize societies, it is also changing classic business and economic paradigms.

Id. at 1-2.

2. See *id.* at 5. In the Report's words, the primary role for government "should be to support and enforce a predictable, minimalist, consistent and simple legal environment for commerce." *Id.* This recognizes that "government agreements may prove necessary to facilitate electronic commerce and protect consumers." *Id.*

conception of Internet governance as decentralized, the Report argues that any regulation necessary to achieve its goals should be implemented in a decentralized rather than top-down manner.³

In the area of law traditionally viewed as commercial or contract law,⁴ the Report generally supports the freedom of parties to contract as they see fit.⁵ The government's role is primarily to supply a set of background rules against which parties may contract.⁶ The Report also notes that the fluid, borderless nature of the medium requires consistent global rules on contract formation and enforcement to encourage parties to enter into electronic commerce transactions.⁷ Simply put, parties will be less willing to conduct transactions electronically if they are uncertain about the governing law or its contents. Thus, a uniform, readily understood legal environment is essential for fostering electronic commercial transactions.

The high-level policy objectives and underlying principles of the Magaziner Report seem to follow quite sensibly from its vision of the Internet as a low-cost medium. The devil, as always, has been in the details of implementing those high-level objectives. This implementation is still in its infancy and is proceeding on a national as well as international stage. While policymakers continue to avow fidelity to the objectives of the Report, the legislative drafting process codifying those objectives has often been characterized by discord.

The following briefly summarizes some of the major efforts to codify aspects of electronic commerce law and highlights various obstacles that may block uniformity. It then examines an example of such legislation—the Uniform Computer Information Transactions Act in its earlier form as

3. *See id.* (“[G]overnments should establish a predictable and simple legal environment based on a decentralized, contractual model of law rather than one based on top-down regulation.”).

4. Note that the overriding principles and policies of the Report apply not only to contract law but also to all facets of electronic commerce including, for example, such diverse concerns as privacy, taxation and maintenance of an adequate telecommunications infrastructure.

5. *See* FRAMEWORK, *supra* note 1, at 10 (“In general, parties should be able to do business with each other on the Internet under whatever terms and conditions they agree upon”).

6. *See id.* (arguing that the government should “support the development of ... [a] legal framework that recognizes, facilitates, and enforces electronic transactions worldwide”).

7. *See id.* at 2 (noting that Internet commerce occurs globally and that for its potential to be realized, “governments must adopt a non-regulatory, market-oriented approach to electronic commerce, one that facilitates the emergence of a transparent and predictable legal environment to support global business and commerce”).

proposed Article 2B of the Uniform Commercial Code (“U.C.C.”).⁸ Finally, this paper considers how the experience of the original drafters of the U.C.C. informs the controversy surrounding Article 2B, and attempts to draw some lessons from that history to help policymakers as they continue to implement the Magaziner Report’s recommendations.

II. CODIFICATION EFFORTS

Certainly, an initial difficulty in codifying principles of electronic commerce is in defining the term “electronic commerce” in a particular context. For example, it could be defined expansively as all transactions in which electronic means are used in some manner however inconsequential, or, more narrowly, as a transaction performed entirely electronically. Between these two extremes are a range of transactions to which the term might refer. The lack of any uniform definition has perhaps contributed to the piecemeal efforts to codify particular aspects of electronic commerce that characterize today’s legal landscape.

For example, before the Magaziner Report was even issued, the United Nations Commission on International Trade Law (“UNCITRAL”) had adopted a Model Law on Electronic Commerce (“Model Law”).⁹ The intent of the law was to facilitate electronic commerce by providing for essentially equivalent treatment of electronic and paper records.¹⁰ The

8. As this article went to press the National Conference of Commissioners on Uniform State Laws (“NCCUSL”) and the American Law Institute (“ALI”) announced that the rules set forth in Article 2B will henceforth be promulgated as part of a stand-alone act—the Uniform Computer Information Transactions Act—rather than as part of the U.C.C. See NCCUSL & ALI, *NCCUSL to Promulgate Freestanding Uniform Computer Transactions Act: ALI and NCCUSL Announce that Legal Rules for Computer Information Will Not be Part of U.C.C.* (visited Apr. 14, 1999) <<http://www.nccusl.org/pressrel/2brel.html>>. This article, however, will continue to refer to Article 2B because there is no indication that the new uniform act will change the substance of the proposed rules and no version of the uniform act was available at the time this article went to press. Moreover, the lessons that can be learned from the U.C.C.’s drafters still apply regardless of the manner of promulgation of rules for transactions in information.

9. *Model Law on Electronic Commerce of the United Nations Commission on International Trade Law*, G.A. Res. 162, U.N. GAOR, 51st Sess., Annex, Agenda Item 148, U.N. Doc. A/RES/51/162 (1996), available at <<http://www.un.or.at/uncitral/english/texts/electcom/ml-ec.htm>>.

10. *Id.* at Guide to Enactment of the UNCITRAL Model Law on Electronic Commerce, para. 6 (“The objectives of the Model Law, which include enabling or facilitating the use of electronic commerce and providing equal treatment to users of paper-based documentation and to users of computer-based information, are essential for fostering economy and efficiency in international trade.”).

Magaziner Report refers to the Model Law favorably, "support[ing] the adoption of [its] principles ... by all nations as a start to defining an international set of uniform commercial principles for electronic commerce."¹¹ A number of individual states, drafting bodies and countries are considering adopting all or part of the Model Law as they update their legislation for the electronic world.¹²

In the United States, the National Conference of Commissioners on Uniform State Laws ("NCCUSL") has been working on a Uniform Electronic Transactions Act ("UETA") against the backdrop of the UNCITRAL Model Law, as well as state enactments on the use of electronic records and digital signatures. As a promulgator of the U.C.C. and other uniform acts, NCCUSL historically has been influential in enabling low-cost transactions across state borders by encouraging the states to adopt uniform laws drafted by NCCUSL, often with the assistance of other bodies such as the American Law Institute ("ALI"). However, NCCUSL's current UETA effort may suffer from tardiness. A number of states are in the process of adopting or have already adopted laws based on the Model Law.¹³ This approach is consistent with the Magaziner Report's suggestion of a decentralized model of regulation but may prove less desirable than states' adoption of the UETA. Individual state enactments, although generally premised on the principles of the Model Law and UETA, will likely differ in their details, potentially frustrating the uniformity essential to the free flow of electronic commerce.¹⁴

Perhaps because of the potential for lack of uniformity among the states and certainly because of a fear of lack of uniformity internationally, the Clinton Administration recently proposed that UNCITRAL consider preparing an International Convention on Electronic Transactions.¹⁵ The

11. FRAMEWORK, *supra* note 1, at 11.

12. See U.S. GOV'T WORKING GROUP ON ELEC. COMMERCE, FIRST ANNUAL REPORT 13 (1998), available at <<http://www.doc.gov/ecommerce/E-comm.pdf>> (noting that "a number of countries and most U.S. States are using the Model Law as a basis for updating their commercial laws") [hereinafter FIRST ANNUAL REPORT].

13. See *id.*

14. This lack of uniformity of course could be avoided by individual state's accompanying their enactment of the UETA with repeal of these prior acts. This approach has been successful in other areas. See, e.g., U.C.C. § 9-102 (1998) (stating that enactment of Article 9 of the U.C.C. "should be accompanied by the repeal of existing statutes dealing [with security interests in personal property]").

15. See FIRST ANNUAL REPORT, *supra* note 12, at 15 (noting that "[a] few governments ... are establishing detailed rules for electronic authentication, which the United States considers to be premature, burdensome or unnecessary").

Convention would eliminate legal barriers to electronic transactions by treating electronic messages as paper equivalents and would provide a framework for authentication rules. In addition, it would synthesize the Model Law and new provisions in pursuit of a coherent whole.¹⁶ In its proposal, the Administration recognized that although different nation-states likely will have varying approaches to authentication, an internationally accepted framework may assure contracting parties that their transactions will be respected despite differing laws at the local level.¹⁷ Although acknowledging that a Convention might be desirable, UNCITRAL has chosen instead to focus on drafting uniform rules regarding digital and electronic signatures to supplement the already existing Model Law.¹⁸

In late 1998, the European Commission ("Commission") presented a Proposal for a European Parliament and Council Directive on Certain Legal Aspects of Electronic Commerce in the International Market ("Commission Proposal").¹⁹ The Commission Proposal is intended to implement some of the objectives announced by the Commission in its 1997 European Initiative on Electronic Commerce.²⁰ That document, although published before the Magaziner Report, bears some similarities to the latter report, including an emphasis on market-driven solutions and a basic belief that the law should encourage global electronic commerce.²¹ In the new Commission Proposal, the Commission has included a provision that directs member states to:

ensure that their legislation allows contracts to be concluded electronically. Member States shall in particular ensure that the legal requirements applicable to the contractual process neither prevent the effective use of electronic contracts nor result in such

16. *See id.* (indicating that the Convention should use provisions of the Model Law to remove barriers to electronic transactions and couple such provisions with an approach to authentication).

17. *See id.*

18. *Proposal by the United States of America*, U.N. GAOR Comm'n on Int'l Trade Law, 33rd Sess., Note by the Secretariat, at 1, U.N. Doc. A/CN.9/WG.IV/WP.77 (1998) available at <http://www.un.or.at/uncitral/english/sessions/wg_ec/wp-77.htm> (stating that the Committee did not wish to be distracted from focusing on digital signature rules but leaving open the possibility of a later Convention).

19. COM(98)586 final [hereinafter Commission Proposal].

20. COM(97)157.

21. *See id.* at 14 (setting forth principles for regulation of electronic commerce including "No regulation for regulation's sake"); *see also id.* at 4 (noting opportunities for Europe to participate in the global technological revolution).

contracts being deprived of legal effect and validity on account of their having been made electronically.²²

This proposal is consistent with the approaches of UNCITRAL and the UETA.

III. OBSTACLES TO UNIFORMITY

The movement to treat electronic communications as paper equivalents thus seems to be well underway. UNCITRAL has provided a model that much of the world is using to fashion its own legislation. The hope is that the overriding principle of accepting electronic communications will remain consistent across jurisdictions, rendering differing details of individual national or state enactments less likely to frustrate global electronic commerce.

However, it is far from clear whether or not this is occurring. It is relatively uncontroversial in these days of advanced technology that parties should be able to form contracts electronically. But there is widespread disagreement on when an electronic message may be attributed to a person and what legal effect such attribution should have. UNCITRAL, the European Union, the Organisation for Economic Co-operation and Development ("OECD"), the individual states of the United States, NCCUSL, and the American Bar Association are just some of the organizations working on the authentication problem. While these bodies often confer with each other, exchanging ideas and building on each other's work, the coordination among them is incomplete. For example, a number of states are in the process of adopting or have already adopted digital signature laws.²³ These laws lack uniform definitions and vary in their approach to a number of issues.²⁴

There is a risk then that the Magaziner Report's approach of seeking international agreement on a broad framework with details to be implemented at the local level may not achieve the desired uniformity. The need

22. Commission Proposal, *supra* note 19, at Art. 9.

23. For a collection of state enactments, see *Summary of Electronic Commerce and Digital Signature Legislation* (last modified Mar. 8, 1999) <http://www.mcb.com/ds_sum.html>.

24. See Thomas J. Smedinghoff, *Overview of State Electronic and Digital Signature Legislation*, Glasser Legal Works (1988), available in Westlaw, TP-ALL Library, EL-CEC GLASS-CLE 407 (noting that among the states there is "little consensus" on how to approach digital signature issues).

for high-level agreement, not just on overarching principles, but also on some details, is becoming apparent.

The Administration's proposal for an International Convention seems to recognize this.²⁵ As different organizations attack facets of electronic commerce in a piecemeal fashion, the risk of a lack of uniformity increases. An International Convention addressing the issues of removing barriers to electronic commerce and agreeing on authentication principles could provide guidelines within which individual states could act without compromising uniformity. As Professor Amelia Boss notes, the law's focus is shifting toward a symbiosis between national and international codification efforts for a number of reasons, including (i) the lack of any existing body of law governing electronic commerce, (ii) the globalization of such commerce, and (iii) the number of nations addressing electronic commerce issues at the same time.²⁶ Old paradigms in which law was drafted domestically and later harmonized with international efforts are giving way to a new paradigm of coordination geared toward creating legal systems that are unified in their approach if not in the details of their rules.²⁷

This suggests that the Clinton Administration's approach should continue to evolve. It has largely been successful in obtaining global agreement on the broad principles outlined in the Magaziner Report. It should now turn its attention both to process matters and to the details of substantive rules. The Clinton Administration should work to ensure that coordination across countries and within the United States occurs, resulting in rules that are consistent with the Magaziner Report's philosophy. This is not likely to be an easy task as nations and states may be reluctant to cede sovereignty to a coordinated effort. Additionally, cultural differences suggest that an overarching consensus will never be reached; and certainly, no consensus will be reached on the minutiae. The key is to work to implement framework principles with enough detail to allow electronic com-

25. See *supra* notes 14-17 and accompanying text.

26. See Amelia H. Boss, *Electronic Commerce and the Symbiotic Relationship Between International and Domestic Law Reform*, 72 TUL. L. REV. 1931, 1943-46 (1998) (identifying a new paradigm for lawmaking in electronic commerce and stating, "A number of identifiable factors have contributed to the symbiosis that exists between international and domestic legal developments in the area of electronic commerce. The relative state of development of law, the timing of the processes, the globalization of commerce and the evolution of new commercial practices are all contributing factors").

27. See *id.* at 1943 ("With the advent of electronic commercial practices, ... the focus shifts from harmonization to coordination, from efforts to bring disparate legal systems together to efforts to create legal systems that are unified in their approach.").

merce to proceed in an environment of legal certainty while not depriving local jurisdictions of the ability to enforce local values through their own laws.

The difficulty in achieving this balance is already apparent in the debates over U.C.C. Article 2B. The high-level policies of the Magaziner Report do not provide much assistance in determining how—or if—to adapt common law contract and U.C.C. rules to a digital environment. There are a multiplicity of contexts in which the law will have to determine the validity of online contract formation and the terms of that contract. For example, online contracts may be for the purchase of tangible goods, intangible information, access to information, or performance of services, just to name a few. Moreover, these contracts and their subject matter may be standard form or highly individualized and between two businesses, a business and a consumer, or two consumers. How—and, more particularly, should—legislators at the state, national or international level draft a law or set of laws to deal with such diverse electronic transactions under current or future technology?

The United States' answer has been to attack aspects of the problem under different legal rubrics. Historically, contracts have been governed primarily by state law. Transactions in goods have been governed by Article 2 of the U.C.C. while contracts for services and information have been governed by common law contract rules. Additionally, contracts for information have always been subject to federal intellectual property laws that preempt state contract law in the case of conflict.²⁸ As electronic commerce has grown, both the individual state governments and the federal government have recognized the need to adjust the relevant law to account for the special concerns it raises.

In the contract law area, NCCUSL's proposed revision of Article 2 would apply to electronic purchases of goods while its proposed Article 2B would cover computer information transactions. Computer information transactions include licenses and other contracts involving software, other information, or access to such information.²⁹ Article 2B applies to these

28. See U.C.C. § 2B-105, Reporter's Notes 2-3 (Dec. 1998 Draft) (noting that federal intellectual property law may preempt state contract law).

29. See *id.* § 2B-102(a)(9). This provision states:

Computer information transaction means a license or other contract whose subject matter is (i) the creation or development of, including the transformation of information into, computer information or (ii) to provide access to, acquire, transfer, use, license, modify, or distribute computer information. The term does not include a contract for distri-

transactions regardless of whether they are entered into on- or off-line or whether the information is delivered electronically or through more conventional means. The revision of Article 2 is scheduled for completion in 1999.³⁰ NCCUSL originally also planned to hold a final vote on Article 2B in 1999.³¹ However, the ALI, citing a number of concerns with the then-current draft, indicated that it would not vote on Article 2B in 1999 and might not even discuss it at its annual meeting.³² In a late-breaking development, on April 7, 1999, the ALI and NCCUSL announced that the rules set forth in Article 2B would no longer be presented as part of the U.C.C.; rather they would form an entirely new and separate enactment—The Uniform Computer Information Transactions Act (“UCITA”).³³ The ALI will not vote on this Act but NCCUSL plans to introduce it to the states for enactment in Fall 1999.³⁴ There is no indication that NCCUSL plans to change the substance of Article 2B’s rules as they are incorporated in the UCITA. Thus, the debate over Article 2B continues to be relevant.

The ALI is not the only group that has expressed reluctance to present Article 2B to the states for enactment. Legal academics, consumer and government groups, and certain industries have voiced numerous objections to Article 2B over the course of its drafting.³⁵ A detailed analysis of Article 2B’s particular provisions and objections to them is beyond the scope of this paper. However, a broad overview of the critical literature indicates that objections run the gamut from suggesting that Article 2B’s

tribution of information in print form ... or to create information for the purpose of distribution in print even if the information provided for distribution pursuant to the contract is delivered in electronic form.

Id.

30. See Kathleen Patchel, *The Uniform Commercial Code Survey Part 1: Introduction*, 53 BUS. LAW 1457 (1998) (stating that revised Article 2 and 2B are “slated for final approval by the ALI and NCCUSL in 1999”).

31. See *id.*

32. See ALI *Concerns About Scope, Public Policy Delay Vote on New Software Licensing Law*, ELECTRONIC COM. & L. REP. (BNA) NO. 48, at 1424-25 (Dec. 23, 1998).

33. See *supra* note 8.

34. See *id.*

35. See, e.g., Symposium, *Intellectual Property and Contract Law in the Information Age: The Impact of Article 2B of the Uniform Commercial Code on the Future of Transactions in Information and Electronic Commerce*, 13 BERKELEY TECH L.J. 809 (1998) (devoting an entire issue of the law journal to articles largely critical of Article 2B); Symposium, *Intellectual Property and Contract Law for the Information Age: The Impact of Article 2B of the Uniform Commercial Code on the Future of Informational and Commerce*, 87 CALIF. L. REV. 1 (1999) (same); NCCUSL, *U.C.C. Article 2B Licenses* (visited Mar. 17, 1999) <<http://www.law.upenn.edu/library/ulc/ulc.htm#ucc2b>> (referencing drafts of Article 2B and correspondence revealing controversial issues).

proposed scope is overbroad to challenging its provisions on standard form contracts and its alleged inadequate attention to public policy concerns.³⁶

The sheer volume of commentary criticizing and defending Article 2B illustrates its importance. It is the first effort to codify the law on transactions in information. As such, it is likely to provide a model not only for other countries individually, but also for international movements to unify laws applicable to cross-border transactions. Moreover, as the law that establishes substantive rules for contract formation and enforcement, it will likely prove the linchpin in the effort to encourage global electronic commerce. The already occurring, universal acceptance of the principle of electronic contracting is a necessary first step to facilitating electronic commerce. However, such commerce, particularly in the information transactions that Article 2B would regulate, is unlikely to reach its full potential in the absence of a legal environment that sets forth coherent, consistent rules.

This makes the acrimony over Article 2B all the more troubling. Commentators widely accept the need for uniform principles, particularly to facilitate online delivery of information.³⁷ Why then is Article 2B faltering? The answer lies in an appreciation both of history and of the nature of the transactions Article 2B is attempting to regulate. A brief overview of the U.C.C.'s history considered in light of current legislative efforts helps to explain why Article 2B has been so controversial and also offers some suggestions for how to arrive at a more satisfactory proposal.

IV. THE LEGACY OF LLEWELLYN: LESSONS FOR THE DRAFTERS OF ARTICLE 2B AND OTHER LAWMAKERS

A. A Brief History of Article 2

Current Article 2 of the U.C.C. was subject to a crucible of debate similar to that now engulfing Article 2B. Like Article 2B, Article 2's drafting process proceeded over a number of years, with the draft being revised to reflect comments received from diverse groups.³⁸ The goals of Article 2's drafters were similar to those of Article 2B's and anticipated by

36. *See id.*

37. *See, e.g.,* FRAMEWORK, *supra* note 1, at 2 (discussing online provision of services and information).

38. *See* Robert Braucher, *The Legislative History of the Uniform Commercial Code*, 58 COLUM. L. REV. 798, 799-804 (1958) (describing the iterative process of presenting drafts, receiving comments and criticism, and redrafting).

50 years the policy recommendations that the Magaziner Report would make with respect to electronic commerce. As stated in the statute,

Underlying purposes and policies of [the U.C.C.] are

- (a) to simplify, clarify and modernize the law governing commercial transactions;
- (b) to permit the continued expansion of commercial practices through custom, usage and agreement of the parties;
- (c) to make uniform the law among the various jurisdictions.³⁹

The drafters sought to make commercial contracting more efficient, reducing costs by eschewing reliance on formalistic and legalistic common law contract rules, and instead giving effect to the parties' agreement as revealed in the particular transaction. Generally, the drafters' philosophy and particularly that of Article 2 Reporter Karl Llewellyn was labeled "legal realism."⁴⁰ Legal realists believe that law can—and should—be revealed by the practices of parties actually engaged in commercial transactions.⁴¹ In translating this philosophy into law, the drafters intended courts to be informed by a "situation sense" and to interpret the parties' agreement in all of the circumstances of the transaction including usage of trade, course of dealing and course of performance.⁴² This orientation reflected the drafters' belief that the parties are best suited to determine their agreement and the law should honor that agreement. Thus, the principle of freedom of contract is enshrined in Article 2 and reflected in its drafting style.⁴³ Under Article 2, parties can make their own bargain by contracting

39. U.C.C. § 1-102(2) (1998). This statement of purpose, set forth in Article 1, applies to each of the following articles, including Article 2 on Sales. See U.C.C. § 1-102, cmt. 1 (1998).

40. See generally WILLIAM TWINING, KARL LLEWELLYN AND THE REALIST MOVEMENT 302-40 (1973) (describing the philosophy and Llewellyn's adherence to it in drafting the U.C.C.).

41. See Maureen A. O'Rourke, *Rethinking Remedies at the Intersection of Intellectual Property and Contract: Toward a Unified Body of Law*, 82 IOWA L. REV 1137, 1153 n.70 (collecting authorities to this effect).

42. See *id.* at 1153-54 n.70.

43. See U.C.C. § 1-102(3) & cmt. 3 (1998) (stating in the statutory text that "[t]he effect of provisions of this Act may be varied by agreement" and explaining in the accompanying comment that "freedom of contract is a principle of the Code").

around most of the statutory rules or simply opt to rely on these default rules as the terms of their contract.⁴⁴

Although not without its problems, Article 2 has been widely regarded as a success, particularly in business-to-business transactions.⁴⁵ It has been adopted in most United States jurisdictions,⁴⁶ offering a measure of national uniformity that has facilitated the growth of a national economy. It would seem then that Article 2B could do much worse than to proceed along the same lines as Article 2.

However, it is by no means clear that the approach which informed Article 2 is appropriate in drafting a law to govern computer information transactions. First, the scope of the two articles is much different. Article 2 governs transactions in goods, defined as "all things ... which are movable at the time of identification to the contract for sale."⁴⁷ Whether standardized or specially manufactured, goods are tangible. The product is physically defined, and the buyer obtains all rights to that physical manifestation on completion of a sale. The same rules that govern the sale of standardized commodities like shampoo or canned corn also are effective to govern the sale of, for example, sophisticated airplane parts that are designed specially for a particular buyer. The fit may be imprecise but it is workable. In attempting to decrease the costs of commercial transactions, Article 2's philosophy was to adopt as the default those rules to which most parties would agree.⁴⁸ Thus, generic contracting parties could save costs by relying on those default rules while parties with special needs could contract around some or all of them as appropriate.

44. *See id.*

45. Commentators as well respected as James J. White and Robert S. Summers have called the U.C.C. "[t]he most spectacular success story in the history of American law." Fred H. Miller, *The Uniform Commercial Code: Will the Experiment Continue?*, 43 *MERCER L. REV.* 799, 808 (1992) (citing JAMES J. WHITE & ROBERT S. SUMMERS, *UNIFORM COMMERCIAL CODE* 3 (3d ed. 1988)). *But see* Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms*, 144 *U. PA. L. REV.* 1765, 1821 (1996) (contending that the U.C.C.'s approach to defining the agreement has some undesirable effects on merchant transactions).

46. *See* ALAN SCHWARTZ & ROBERT E. SCOTT, *COMMERCIAL TRANSACTIONS-PRINCIPLES AND POLICIES* 2 (2d ed. 1991) (noting that at least parts of the U.C.C. have been adopted in all 50 states).

47. U.C.C. § 2-105(1) (1998).

48. *See* SCHWARTZ & SCOTT, *supra* note 46, at 21 (explaining the economic theory informing the choice of default rules).

B. Article 2B and Information Transactions

The computer information transactions governed by Article 2B are different. As Article 2B itself notes, the product it governs is intangible, not defined physically, but rather by the bundle of rights granted under a license.⁴⁹ The rules that govern a license of software protected by a copyright or patent may not be suitable for a license of information protected by a trade secret, or for agreements granting access to information. The “one size fits all” approach of Article 2 simply may not work.

Article 2B’s drafters have recognized this, setting forth different rules depending on the nature of the transaction. For example, the rules for access contracts differ from the rules for software licenses.⁵⁰ Article 2B is an umbrella statute covering diverse transactions. This makes it quite unlike Article 2 which deals with the fairly unitary transaction of the sale of goods. The net result is that Article 2B, in trying to be all things to all types of computer information transactions, has reached a level of complexity that is at odds with both the U.C.C.’s and the Magaziner Report’s goals of simplicity and clarification.

This complexity becomes all the more apparent when one compares Article 2B to the Commission Proposal. The Commission Proposal applies to “Information Society services,” an overarching concept that includes “all services normally provided against remuneration, at a distance by electronic means,” including the online sale of goods and services and the license of information.⁵¹ Thus, the unifying factor is the means of contracting rather than the subject matter of the contract. This approach has some intuitive appeal and may be simpler as well as more coherent than Article 2B’s scope definition.⁵² This suggests that Article 2B might benefit from some synthesis with the Commission Proposal—a competition between the two approaches might ultimately lead to a “better” law than if each were drafted without reference to the other.

49. See U.C.C. § 2B-102, Reporter’s Note 38 (Dec. 1998 Draft) (“Scope provisions in a license define the product.”).

50. See *id.* § 2B-615 (setting forth the rules for access contracts).

51. See Commission Proposal, *supra* note 19, at 14-15.

52. Note, however, that Article 2B’s drafters considered a number of approaches including a “hub and spoke” approach under which the “hub would consist of general contract principles and the scope rules for particular types of transactions like sales, licenses or leases.” See Raymond T. Nimmer, *Intangibles Contracts: Thoughts of Hubs, Spokes, and Reinvigorating Article 2*, 35 WM. & MARY L. REV. 1337, 1340-1341 (1994). However, the eventual result was to draft Article 2B as a stand-alone article containing all of the substantive contract rules relating to computer information transactions.

C. Article 2B and Intellectual Property Law

The intricacy of Article 2B is also caused by its need to account for federal intellectual property law and the practices of knowledge-based industries, a concern that Article 2 never had to face. Some of the main criticisms leveled against Article 2B focus on its alleged failure to adequately accommodate federal intellectual property policy considerations.⁵³ Critics argue that Article 2B is attempting to preempt federal law through the mechanism of state contract law, a clearly impermissible approach.⁵⁴

Article 2B contains a provision stating the obvious proposition that federal law preempts state law as well as a rule giving courts' broad freedom to reform contracts containing terms that violate "fundamental public policy."⁵⁵ The "public policy" exception is a compromise provision intended to redress some parties' objections that Article 2B did not go far enough to enshrine the federally-struck balance between the rights of creators to protect their information and the rights of the public to use it.⁵⁶ The problem for Article 2B's drafters is that the very debate that resulted in this provision itself illustrates that there is no consensus on how that balance has been struck.

This graphically illustrates the internal conflicts of the U.C.C. Its goals of flexibility and uniformity can be at war with each other. This conflict historically has been papered over in the Article 2 context. Article 2 uses a number of flexible concepts like good faith and commercial reasonableness in its statutory wording.⁵⁷ The states enact uniform wording but interpretations vary. Good faith in Maine may not be the same thing as good faith in Texas which may differ from good faith in California. Results may vary across jurisdictions despite the fact that the different locales are in-

53. See Raymond T. Nimmer, *Breaking Barriers: The Relation Between Contract and Intellectual Property Law*, 13 BERKELEY TECH. L.J. 827 (1998) (defending Article 2B's treatment of intellectual property rights).

54. Cf. Jessica Litman, *The Tales that Article 2B Tells*, 13 BERKELEY TECH. L.J. 931, 941 (1998) (contending that Article 2B encourages licensors to assert by contract rights which are not afforded under copyright law).

55. U.C.C. § 2B-105(a)-(b) (Dec. 1998 Draft).

56. The "McManis motion" brought forward by Professor Charles McManis began the debate which culminated in § 2B-105. See J. Thomas Warlick IV, *A Wolf in Sheep's Clothing? Informational Licensing and De Facto Copyright Legislation in U.C.C. 2B*, 45 J. COPYRIGHT SOC'Y U.S.A. 158, 165-166 (1997) (describing the evolution of § 2B-105).

57. See, e.g., U.C.C. § 2-706(1) (1998) (setting forth the calculation of an aggrieved seller's remedy "where the resale is conducted in good faith and in a commercially reasonable manner"); see also *id.* § 1-203 ("Every contract or duty within the Act imposes an obligation of good faith in its performance or enforcement.").

terpreting exactly the same statutory language. This flexible approach to law left room for local values to be enforced and for the statute's meaning to change with time, obviating the need for continual revisions to the statutory text, but sacrificing some degree of uniformity in the process.

Article 2B's drafters are continuing this approach in striking a balance with federal intellectual property law. The flexible language that they have proposed makes it more likely that individual states will enact the statute and enables varying interpretations across these states. The point of contention is that there is a substantive difference between allowing states to have varying interpretations of contract law concepts like good faith and commercial reasonableness and allowing states to have varying interpretations of the intellectual property rights granted under federal law. A more explicit delineation of which intellectual property rights and obligations may be contracted away and which may not, is more likely to ensure that individual states respect the federal balance.

The Magaziner Report is silent on this debate. It does not specifically address the relationship between intellectual property and contract law. It does separately advocate freedom of contract and sufficient protection for intellectual property rights to provide incentives for creators to commercialize their works electronically.⁵⁸ However, it also notes that nations may want to allow for exceptions to such rights like fair use.⁵⁹ It does not state whether a party can give up its fair use rights by contract.

Its framework approach, global orientation, and belief in freedom of contract suggest that it might advocate the drafters' flexible approach. However, this is by no means clear and it is likely that other groups, such as the European Union, would disagree. In the last several years, the European Union has begun to make more explicit what provisions of its Directives—including those related to intellectual property—may not be contracted around.⁶⁰ It has also frequently conditioned protection for the information of non-Union entities on principles of national treatment.⁶¹

58. See *supra* note 5; see also FRAMEWORK, *supra* note 1, at 12 (stating that providers will not make works available electronically if they believe that their intellectual property will be stolen).

59. See FRAMEWORK, *supra* note 1, at 13.

60. For example, the European Directive on the Legal Protection of Computer Programs provides that parties may not contract around the limited decompilation right granted under the Directive. See Directive 91/250/EEC of the Council of 14 May 1991 on the Legal Protection of Computer Programs, art. 6, 1991 O.J. (L 122) 42.

61. For example, the EC Database Directive conditions certain database protections on "third countries offer[ing] comparable protection to databases produced by nationals of a Member State." Directive 96/9/EC of the European Parliament and of the Council of

These two European trends indicate that clarifying domestic law and harmonizing it with other nations' laws has assumed added importance. The opposition to Article 2B has identified an important issue, and the Administration's role now should be to consider how best to address it—in Article 2B, in intellectual property statutes, or in an international forum.

D. Article 2B, the “Agreement,” and Consumer Protection

The difference in subject matter that has made Article 2B a much more complex endeavor than Article 2 is exacerbated by the fast-moving nature of the industries that would be subject to its provisions. Under both Articles 2 and 2B, agreement is defined as “the bargain of the parties in fact as found in their language or by implication from other circumstances including course of dealing or usage of trade or course of performance.”⁶² Article 2 was drafted against the backdrop of hundreds of years of commercial law development from Roman law through the Middle Ages to the English Sale of Goods Act of 1893 and finally to state enactments in the United States.⁶³ It made sense for the statute to talk about usage of trade because longstanding practices had grown up that were consistently observed within particular industries. These practices formed the backdrop against which parties contracted. Incorporating such usage into an agreement was truly a cost-saving device. The parties expected such terms to apply and were freed from memorializing them in the contract by the U.C.C.'s default incorporation of them into the “agreement.”

In many industries which would be governed by Article 2B, it is premature to refer to a usage of trade. Customs are rapidly evolving, and deference to a particular norm at a particular time may not be appropriate. In some ways this offers a great opportunity to the drafters of Article 2B that Article 2's drafters lacked. Article 2B could channel norms toward the development of desirable practices.

For example, Article 2B sanctions “pay now, terms later” contracts or shrinkwraps as well as online contracts in which the customer manifests assent by clicking on an icon.⁶⁴ However, the governing contractual terms

11 March 1996 on the Legal Protection of Databases, preamble para. 56, 1996 O.J. (L 77) 20.

62. U.C.C. § 1-201(3) (1998).

63. See ROBERT BRAUCHER & ROBERT A. RIEGERT, INTRODUCTION TO COMMERCIAL TRANSACTIONS 19-21 (1977) (outlining pre-U.C.C. sources of law.).

64. See, e.g., U.C.C. § 2B-207, Reporter's Note 3 (Dec. 1998 Draft) (discussing “layered contracting” in which terms follow performance and how 2B sanctions such contracts).

themselves may not be displayed, but simply available to the customer if it chooses to "click-through" to another page to read them. The assent is effective regardless of whether the customer actually clicks through or not.

Under Article 2, the use of terms that were provided after payment could enable a mass market. Often in faceless transactions, it is difficult for the seller to communicate terms to the buyer or to incorporate all of them visibly on the exterior of the packaging to be read before purchase. Thus, it might be reasonable to have a legal rule that allowed the seller to put terms inside the box, but gave the buyer a right to reject those terms once he saw them.

But should such a rule apply to online transactions? The transaction is still faceless but the difficulty of communication no longer applies. The seller can easily provide the terms on the screen for the buyer's examination. Moreover, the seller can easily, through technology, assure that the buyer cannot assent to the terms without those terms having first been presented on the screen. Should it be enough for the seller to have the terms available someplace on the site or should the seller be required to place the terms before the buyer prior to assent?

In other words, should not the fact that Articles 2 and 2B are dealing with different markets matter?⁶⁵ Rules that facilitate mass markets in tangible goods may not be ideal for the electronic mass market. Already available technology makes it possible for the electronic market to inform customers about the terms of a deal much more easily than in the tangible world. But the current usage of trade (to the extent one exists) may be to place the terms on another part of a website rather than on the customer's screen prior to assent. Should the law enshrine such usage of trade or encourage adoption of another one? Certainly, many believe that the electronic market can be very competitive, offering greater choice to customers.⁶⁶ Would it not be consistent with the Report to draft legal rules that enhance the market's efficiency by providing more information to customers, correcting for information asymmetries that might otherwise exist and distort market performance?

65. Keep in mind, however, that Article 2 would apply to electronic contracts for the sale of goods. It too may need some adjustment to account for the unique ability of the electronic media to allow buyers to view terms prior to purchase. However, a critique of the Revised Article 2 is beyond the scope of this article. The unique capabilities of electronic contracting do though offer further support to the European approach in contrast to that of the United States. See *supra* text accompanying notes 51-52.

66. See Commission Proposal, *supra* note 19, at 6 (stating that electronic commerce offers the potential to provide a wider variety of goods and services at lower prices).

Does it not also matter who the “customer” is? Article 2B’s drafters would do well to learn from Article 2’s history. Article 2 does not contain many consumer protection provisions. Over the years, federal and non-uniform state enactments in the consumer protection area effectively either preempted or modified parts of Article 2,⁶⁷ reflecting emerging views of relevant differences between business-to-business transactions and business-to-consumer transactions. Article 2B has the opportunity to incorporate consumer protection into its text, avoiding the later preemption that characterized Article 2. Certainly, Article 2B’s drafters would argue that they have already done so.⁶⁸ The point, though, is that Article 2 provides a lesson—consumer protection included now might avoid nonuniformity later.

The Magaziner Report is largely silent on consumer protection, simply acknowledging that government agreements may be required to protect consumers. However, the OECD has recently issued a Draft Declaration on Consumer Protection in the Context of Electronic Commerce.⁶⁹ The Declaration emphasizes the need for effective consumer protection and discloses the OECD’s intent to develop guidelines for such protections.⁷⁰ The Administration should continue to work closely with OECD on this topic and consider whether or not Article 2B is the appropriate venue for consumer protection or whether it should be enacted in some other form.

E. Summary: Striking a Balance Between Uniformity and Flexibility

The problem that Article 2B faces on the consumer protection front is the same as it faces with respect to virtually all of the issues it is addressing. It is also the same as the Administration is currently facing in the area of digital signatures. What is the appropriate balance between flexibility and uniformity? How does one arrive at a statute that sets a framework on which there is broad national and international agreement while providing enough detail so that local enactments do not frustrate uniformity?

There is no simple answer to this question. However, the sheer difficulty of the task suggests that the key in enabling global electronic commerce may be in obtaining agreement on jurisdictional issues including

67. One area in which supplementary laws have often been enacted is warranty. See SCHWARTZ & SCOTT, *supra* note 46, at 225 (noting both state and federal enactments in the warranty area).

68. See *e.g.*, *infra* notes 73-76 and accompanying text (setting forth Articles 2B’s choice of law rules that ensure that the consumer retains protection under his state’s law).

69. Doc. DSTI/CP(98)12/REV2 (Oct. 8-9, 1998).

70. See *id.*

what contractual choice of law and choice of forum clauses will be enforceable. If parties can choose the law that will govern their transaction and know that a court will respect their choice, transaction costs should decline. An information supplier would not have to know the minute details of every local law but simply the law which governs its particular transactions. It may then estimate its liability exposure with more certainty, decreasing its costs and increasing its willingness to market to a broader audience.

Unfortunately, matters are not so simple. Any discussion of choice of law provisions must also address what rules various jurisdictions consider mandatory. Mandatory rules may be defined as those that may not be contracted around either explicitly, through a specific contractual clause, or implicitly, through a choice of law provision that selects a jurisdiction which has not adopted a particular rule as the governing law. For example, under the Commission Proposal, "the country in which an online merchant is 'established' has the exclusive authority to regulate its conduct."⁷¹ However, consumer contracts are governed by rules requiring online merchants to comply with the mandatory rules (usually including consumer protection laws) of the consumer's jurisdiction.⁷² Online merchants may thus have to incur the costs of learning the mandatory rules of each jurisdiction from which a consumer may transact business with them over the Internet—an expensive proposition which may decrease their willingness to market online.

Article 2B, even more than the European Union, explicitly adopts this "know all laws" approach. Article 2B's default choice of law is the "jurisdiction in which the licensor is located when the agreement is made."⁷³ However, a "consumer transaction that requires delivery of a copy on a physical medium is governed by the law of the jurisdiction in which the copy is or should have been delivered to the consumer."⁷⁴ Moreover, while the parties are free to contract around these provisions, "in a consumer transaction the [parties'] choice is not enforceable to the extent it

71. See Mathew S. Yeo & Marco Berliri, *Conflict Looms Over Choice of Law in Internet Transactions*, ELECTRONIC COM & L. REP. (BNA) NO. 4, at 87 (Jan. 27, 1999) (describing the European approach and noting that it is based on the "country of origin" principle).

72. See *id.* at 86-88 (explaining how the Rome Convention, Distance Selling Directive and Electronic Commerce Directive lead to the conclusion that consumers receive the benefit of the protection of "mandatory rules" of their own jurisdiction).

73. U.C.C. § 2B-107(b)(1) (Dec. 1998 Draft).

74. *Id.* § 2B-107(b)(2).

would vary a rule that may not be varied by agreement under the law of the jurisdiction whose law would apply in the absence of the agreement.”⁷⁵ The drafters recognized that this approach may be expensive for online merchants but argued that the consumer protection choices of individual states should be recognized.⁷⁶

Under both the European and the Article 2B approaches, merchants may be less willing to sell online if they will be subject to widely varying consumer protection laws. If they do sell electronically, they will have an incentive to “zone” customers by country of origin, offering terms complying with the particular country’s law. Smaller firms that cannot incur the costs of knowing the laws of every country may adopt a high-tech version of exclusionary zoning, refusing to sell to customers residing in countries with laws the merchant either does not know or does not like.⁷⁷

75. *Id.* § 2B-107(a).

76. *Id.* § 2B-107, Reporter’s Note 4. (“This rule imposes significant costs on Internet commerce, but this article adopts the view that the fundamental policy of freedom of contract should be varied to preserve consumer rules when individual states, having addressed that cost separately, determine that the applicable rule is of a mandatory, non-waivable nature.”).

77. There may be another alternative for small firms. These firms could rationally choose to remain ignorant of most of the world’s laws and still sell globally, so long as they knew the law of those countries where the majority of their customers are likely to reside. Even on the global medium of the Internet, a firm is likely to be able to estimate in what jurisdictions the majority of its customers are likely to reside. The firm can then protect itself by learning the law of these jurisdictions, a much smaller subset of rules than the laws of all countries. For example, a small U.S. firm is likely to have or obtain some sense of U.S. federal and state laws. Before marketing on the Internet, it might also seek to understand the law of the European Union. Depending on the firm’s product, these two markets may be likely to drive most of the firm’s sales. The firm could then quite rationally choose to remain ignorant of the rest of the world’s laws yet still sell globally. This firm would compare the cost of understanding those laws and adjusting the terms of sale for citizens of other countries through technological means with the likely loss it would incur defending against suits brought by citizens outside its target markets. The likely loss would include the sum of the costs of both successful and unsuccessful consumer suits. The cost of successful suits would be a function of the probability that a customer from a remote country would purchase from the firm and the probability that a consumer would sue and succeed, multiplied by the costs to the firm, including damages. The cost of an unsuccessful suit would be a function of the probability of purchase and the probability of a non-successful suit multiplied by the costs to the firm of defending against suit. Because both the probabilities and costs may be quite low, the expected loss may also be quite low, far outweighed by the cost of adjustment. Thus, a firm could rationally sell globally while remaining relatively ignorant of other jurisdictions’ laws. This strategy may become less feasible as larger non-U.S. and European Union markets, like China and Africa, become major sources of customers for Internet-based firms.

This situation illustrates the problem that Article 2B and any other legislation that is not international in character may have in achieving uniformity in a global medium. It suggests that the Clinton Administration should adopt a two-pronged approach to global electronic commerce.

The first facet of the strategy would focus on the international harmonization of laws which nation-states regard as mandatory. For example, if the mandatory and varying nature of consumer protection laws should begin to hamper electronic commerce, the Clinton Administration should seek international agreement on a minimum level of consumer protection. Nations would be free to enact greater levels of protection, but merchants would be bound only by the minimum. This does not mean that all merchants would gravitate to that minimum. Many might "opt in" to the laws of nations with a higher degree of consumer protection to signal to their customers that their product possesses qualities superior to those sold under the lower standard.⁷⁸

Admittedly, it would be quite difficult to obtain international agreement on a minimum level of consumer protection.⁷⁹ However, if requiring that the laws of either the merchant's establishment or the consumer's residence govern prevents the Internet from realizing its electronic commerce potential, then this approach would still be preferable. The former would risk a race to the bottom among jurisdictions' consumer protection laws,⁸⁰ while the latter would impose substantial burdens on global electronic commerce.⁸¹

The second prong of the Clinton Administration's approach should be to bring Article 2B issues out into the international arena. As already demonstrated, Article 2B alone cannot guarantee global uniformity. By subjecting Article 2B to the crucible of international debate in conjunction with other efforts like the Commission Proposal, however, the Clinton

78. This is similar to the signaling function that the warranty law serves. Sellers often make quality assurances in excess of those imposed under Article 2 of the U.C.C. in order to signal to buyers that their product is superior. See SCHWARTZ & SCOTT, *supra* note 46, at 106 (explaining the signaling function of warranties).

79. See Yeo & Berliri, *supra* note 71, at 89 (arguing for an international consumer protection harmonization effort while noting that "one can hardly minimize the political and procedural complexities of negotiating a uniform law for Internet-based transactions...").

80. See *id.* at 88 (stating that, while allowing merchants to choose the governing law has the virtue of simplicity, "it poses the obvious problem that unscrupulous online merchants could locate themselves in countries with 'favorable' laws, to the likely detriment of their customers").

81. See *id.* at 89 (rejecting the "mandatory rules" approach as too expensive).

Administration may be better able to realize its goals of facilitating global economic commerce. In addition, substantially better rules may result from the competition among different proposals.

This is not to say that Article 2B is a failure. While its complexity, provisions on manifestation of assent, and broad view of contract formation may be at odds with the Magaziner Report, it contains many provisions consistent with that Report. For example, it is drawn to be technology-neutral,⁸² and to allow for new developments like contract formation and performance through the use of electronic agents.⁸³ Additionally, its drafting process has been characterized by a great deal of private sector involvement.⁸⁴ The authors of the Magaziner Report would approve of all of these traits.

The point is that Article 2B's work should not be abandoned. Rather, it should provide the basis for a national and international discussion of how best to reconcile the sometimes conflicting goals of flexibility and uniformity in the context of an overarching desire to encourage global electronic commerce.

V. CONCLUSION

The Magaziner Report's overall philosophy is alive and well in the United States and other countries. It is the implementation of that philosophy that is proving controversial. The risk is that discussions will disintegrate into nonuniform law that will frustrate the growth of global electronic commerce.

Fortunately, no parties to the debate have an interest in seeing such a breakdown. This should help to make domestic and international discussions more fruitful as the common goal dominates. The Administration should continue its work informed by a sense of history such as that involved in drafting Article 2, and with an understanding of when the unique

82. See, e.g., U.C.C. § 2B-102, Reporter's Note 3 (Dec. 1998 Draft) (noting that Article 2B's approach in the digital signature area is intentionally drawn to be "technologically neutral").

83. See, e.g., *id.* § 2B-204 (providing rules for offer and acceptance by electronic agents).

84. However, some commentators argue that the drafting process is subject to capture by powerful interest groups. See generally Jean Braucher, *The U.C.C. Gets Another Rewrite, Just When You Thought You Really Knew the Uniform Commercial Code, Almost Every Article Is Undergoing Changes in a Major Revision*, 82 A.B.A. J. 66, 68 (1996) (noting that an early draft of Article 2B was criticized as too favorable to software providers).

nature of the Internet requires special rules and when it does not. It should also carefully consider, issue-by-issue and transaction-by-transaction, whether or not international agreement is needed. Moreover, it should consider when parties should be allowed to opt out of legal rules and when those rules should be mandatory. Throughout these inquiries, the symbiosis between domestic and international legal development should continue, hopefully leading to a time when any "Update on the Magaziner Report" can say that its goals—at least in the area of commercial law—have been fully realized.

ELECTRONIC COMMERCE SYMPOSIUM

THE NEW MONEY

By Kerry Lynn Macintosh[†]

ABSTRACT

Professor Macintosh analyzes the strategic advantages and disadvantages of credit cards for buyers and sellers. She concludes that Internet commerce needs electronic money in order to achieve its full potential. Professor Macintosh further reasons that the Internet needs “global electronic currencies” that can serve as universal media of exchange, global units of account, and stable stores of value. However, burdensome laws and regulations could delay, or even preclude, the emergence of electronic money. Professor Macintosh concludes that federal regulations and uniform state laws designed to combat money laundering should not apply to electronic payment products. Also, Congress and the state legislatures should work to repeal outdated laws that stand in the way of electronic money.

TABLE OF CONTENTS

| | | |
|------|---|-----|
| I. | INTRODUCTION | 659 |
| II. | CREDIT CARDS ARE NOT ENOUGH | 661 |
| III. | THE INTERNET NEEDS GLOBAL ELECTRONIC CURRENCIES | 664 |
| IV. | HOW CAN WE FREE UP THE SYSTEM? | 666 |
| | A. Federal regulations..... | 666 |
| | B. Uniform law projects | 669 |
| | C. Outdated federal and state laws | 671 |
| V. | CONCLUSION..... | 673 |

I. INTRODUCTION

Nearly two years ago, the Clinton Administration issued *A Framework for Global Electronic Commerce*.¹ The Framework is one of the most

© 1999 Kerry Lynn Macintosh.

[†] Professor of Law, Santa Clara University School of Law. J.D., 1982, Stanford Law School. I thank Scott Pesetsky, Santa Clara University School of Law, Class of 1999, for his research assistance. I am also grateful to the editorial staff of the *Berkeley Technology Law Journal* for their helpful comments and edits.

1. WILLIAM J. CLINTON & ALBERT GORE, JR., A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE § 3 (1997) available at <<http://www.iitf.nist.gov/eleccomm/ecommm.htm>> (hereinafter FRAMEWORK).

radical political documents of this century—not for what it committed government to do, but for what it committed government *not* to do.

The Framework established five basic principles to guide development of Internet commerce. First, the private sector should lead.² Second, governments should avoid undue restrictions that might distort development of the electronic marketplace.³ Third, government should work to foster a legal environment that is predictable, consistent, and minimalist.⁴ Fourth, governments should recognize that the Internet is unique, and requires new policies.⁵ Fifth, electronic commerce should be facilitated on a global basis.⁶

The Framework identified electronic payment systems as a key element of global electronic commerce.⁷ It recognized that, at this early stage in the development of electronic payment systems, the commercial and technological environment was changing quickly, making it hard to develop timely and appropriate policy.⁸ For these reasons, the Framework concluded, inflexible and highly prescriptive regulations and rules would be inappropriate and potentially harmful.⁹ Instead, electronic payment experiments should be monitored on a case-by-case basis.¹⁰

Since the Framework was released, Internet commerce has increased rapidly in volume, and is projected to be hundreds of billions of dollars by the start of the twenty-first century.¹¹ Because sales cannot go forward without payments, the development of electronic payment systems has emerged as a top priority for innovators and policymakers in the new millennium.

Thus far, credit cards have emerged as the most popular method of payment over the Internet.¹² Consumers send their card numbers over the

2. *See id.* at Principles.

3. *See id.*

4. *See id.*

5. *See id.*

6. *See id.*

7. *See* FRAMEWORK, *supra* note 1, § I.2.

8. *See id.*

9. *See id.*

10. *See id.*

11. *See* U.S. GOV'T WORKING GROUP ON ELEC. COMMERCE, FIRST ANNUAL REPORT 1 (Nov. 1998) [hereinafter FIRST ANNUAL REPORT].

12. *See* Peter Wayner, *Electronic Cash for the Net Fails to Catch On*, N.Y. TIMES ON THE WEB, Nov. 28, 1998 (visited April 17, 1999) <<http://www.nytimes.com/library/tech/98/11/cyber/articles/28cash.html>>.

phone lines, apparently confident that existing encryption protocols are sufficient to protect them against theft and fraud.¹³

Meanwhile, competing electronic payment systems are struggling to survive. DigiCash, Inc. is known as the company that developed eCash, a system for making anonymous electronic payments using digital "coins."¹⁴ However, the idea failed to catch on, and DigiCash petitioned for Chapter 11 reorganization in November 1998.¹⁵ Smart cards have not fared much better.¹⁶ Recently, Citibank and Chase Manhattan ended a smart card pilot program operating in the Upper West Side of Manhattan due to a lukewarm response from the public.¹⁷

In this essay, I consider three questions. First, does global electronic commerce need electronic payment systems other than credit cards? Second, if so, what are the characteristics of these systems? Third, what, if anything, can government do to promote the emergence of the necessary systems?

II. CREDIT CARDS ARE NOT ENOUGH

Internet buyers seem to prefer credit cards to other electronic payment systems that have been made available to them.¹⁸ Why?

One reason may be simple familiarity. Internet commerce is still new and intimidating to many. It is easier for buyers to make purchases on the

13. *See id.*

14. *See* David Einstein, *Day Early, Dollar Short—DigiCash Files Chapter 11*, S.F. CHRON., Nov. 6, 1998, at B1-2. Under the DigiCash system, a customer uses her computer to generate a random serial number. The number serves as a digital "coin," and has a dollar value associated with it. The customer submits the coin to her bank, which adds its digital signature, and debits her account. The customer now holds an electronic bank obligation, which she can transmit anonymously as payment for online goods or services. The merchant who receives the coin contacts the issuing bank to verify that it has not been spent before. If the coin is still good, the merchant deposits it in his own bank. *See* Kerry Lynn Macintosh, *How to Encourage Global Electronic Commerce: The Case for Private Currencies on the Internet*, 11 HARV. J.L. & TECH. 733, 735 n.9 (1998).

15. *See* Einstein, *supra* note 14, at B1.

16. A smart card is a plastic card with an embedded computer chip. The chip is loaded with value. To purchase goods or services, a buyer takes the card to a store equipped with a card-reading terminal. After the sale is complete, the store submits the value to the card issuer for redemption. *See* CONGRESSIONAL BUDGET OFFICE, EMERGING ELECTRONIC METHODS FOR MAKING RETAIL PAYMENTS 9-11 (1996); Macintosh, *supra* note 14, at 734.

17. *See* Saul Hansell, *Got a Dime? Citibank and Chase End Test of Electronic Cash*, N. Y. TIMES, Nov. 4, 1998, at C1.

18. *See* Wayner, *supra* note 12.

Internet when they can use a familiar payment method, like the credit card. As time passes, buyers should become more comfortable with Internet commerce. Innovative payment products, such as smart cards and electronic money, should become more familiar to them.

Even then, however, Internet buyers may continue to prefer credit cards, particularly when making expensive purchases. This is because credit cards offer strategic advantages to buyers in general. Consider these two points:

- 1) Every use of a credit card involves a loan to the buyer. This enables her to buy more than she earns. The loan may even be interest-free, if she pays her account off every month. By contrast, if she holds electronic money, she is, in effect, making an interest-free loan to the company that issues the money.¹⁹
- 2) If a credit card purchase goes sour, a buyer often can avoid loss by asserting her claims or defenses on the purchase against the issuer of the card.²⁰ By contrast, once spent, cash cannot be recovered.²¹ Similarly, a cashier's check—long recognized as a substitute for cash—cannot easily be stopped.²² Electronic money that functions like cash or cashier's checks may face similar constraints.²³

19. Smart cards loaded with value, eCash, and similar products are not legal tender; rather, they represent claims against the issuer. *See* Task Force on Stored-Value Cards, *A Commercial Lawyer's Take on the Electronic Purse: An Analysis of Commercial Law Issues Associated with Stored-Value Cards and Electronic Money*, 52 *Bus. Law.* 653, 670 (1997). In effect, the temporal gap between creation of value and redemption of value involves an extension of credit by the issuer. *See id.* at 664.

20. For example, under federal law, the issuer of a bank credit card or convenience card (like American Express) is subject to claims and defenses arising out of a credit card transaction, if: (1) the cardholder makes a good faith effort to resolve her dispute with the merchant; (2) the amount of the transaction exceeds \$50; and (3) the place where the transaction occurred is in the same state or within 100 miles from the cardholder's residence. *See* 15 U.S.C. § 1666(i) (1998); 2 BARKLEY CLARK & BARBARA CLARK, *THE LAW OF BANK DEPOSITS, COLLECTIONS, AND CREDIT CARDS* ¶ 15.07[2] (rev. ed. 1999). Liability is limited to the amount of credit outstanding on the transaction when the cardholder first notifies the issuer of the claim or defense. *See id.* When the seller has issued the credit card, or is under the control of the issuer, or is a franchised dealer, the cardholder can assert her claims or defenses without regard to dollar amount or geographic location. *See id.*

21. *See, e.g., Miller v. Race*, 97 Eng. Rep. 398 (K.B. 1758) (money cannot be recovered once paid honestly upon a valuable consideration).

22. Issuing banks prefer to honor their cashier's checks in order to protect their own credit reputations. The Uniform Commercial Code reinforces this strong self-interest.

However, buyers are only one side of the coin. For *sellers*, credit cards have the following strategic disadvantages:

- 1) Unlike cash or cash equivalents, credit card charges are subject to percentage fees.²⁴ These charges erode profit margins, particularly on inexpensive goods and services.
- 2) As explained above, a buyer who uses a credit card may refuse to pay the issuer on the grounds that she has a claim or defense arising out of the underlying transaction.²⁵ When this happens, the issuer may pass the loss back to the seller.²⁶
- 3) Enrolling in the credit card system requires establishing a relationship with a depository bank, including the signing of a complex commercial agreement.²⁷

The reason Internet buyers are able to insist on credit card use is because they enjoy a bargaining advantage over online sellers—at least for now. The current success of Internet commerce depends on large numbers of consumers, and therefore they must be coaxed into online purchasing. But this state of affairs is not likely to last long. Electronic commerce will continue its explosive growth. More and more buyers will want to participate. Meanwhile, a larger and more diverse complement of sellers will move online. Consumers will sell goods or services to other consumers. Hobbyists will market digital crafts. Retirees will offer consulting services.

As the cybermarket matures and diversifies, the balance of bargaining power will shift. Internet sellers who offer low-cost products will not want

Banks that wrongfully refuse to honor their own cashier's checks can be sued for expenses, loss of interest, and consequential damages. *See* U.C.C. § 3-411(b) (1995).

23. For example, some lawyers believe courts would treat smart cards as cash, and refuse to permit stop payment. This is likely in systems where value cannot be linked to an individual after payment has been made. *See* Task Force on Stored-Value Cards, *supra* note 19, at 720. To avoid uncertainty, issuers may choose to address such legal problems through contracts or system rules. *See id.*

24. Banks that process credit card slips for sellers charge a percentage fee known as the "discount rate." The discount rate ranges from one-half to seven percent of the amount of the credit card slip. *See* 2 CLARK & CLARK, *supra* note 20, ¶ 15.02.

25. *See supra* note 20 and accompanying text.

26. The issuer may charge the amount of the purchase back to the seller's bank, which, in turn, may charge back against the seller pursuant to a recourse agreement. *See* 2 CLARK & CLARK, *supra* note 20, ¶ 15.02[4][b].

27. For an overview of terms included in bank-merchant agreements, *see id.*

to pay percentage fees. Those who transmit information goods or services electronically may not want to accept the risk that buyers might reverse the charges later. Consumers and others who make only occasional sales may not be willing or able to enroll in the credit card system.

Thus, it is much too early to conclude Internet commerce can—or should—rely primarily on credit cards. Soon, Internet sellers who do not like the cost or risk associated with credit cards will demand money from buyers, just as sellers in real space often do. When that happens, the market will need electronic currencies that can circulate from computer to computer, around the world.²⁸

III. THE INTERNET NEEDS GLOBAL ELECTRONIC CURRENCIES

If global electronic commerce does need additional electronic payment systems, what should the characteristics of those systems be? I have previously argued that the Internet needs “global electronic currencies”—that is, currencies that are privately issued, managed, and denominated.²⁹ Companies that issue such currencies will compete with each other for the business of Internet buyers and sellers.³⁰ Only currencies with stable value and wide acceptance in the marketplace will survive.³¹

Global electronic currencies will benefit Internet commerce in three ways. First, the currencies can serve as universal *media of exchange*. Once a user acquires a global electronic currency, she can enter into transactions around the world without having to pay exchange fees.³² Second, the currencies will provide global *units of account*, enabling buyers and sellers all over the world to understand what goods and services are worth without calculation.³³ Third, and perhaps most importantly, global electronic cur-

28. This is not a technological pipe dream. Mondex Co. possesses technology that permits users to transmit stored value directly from smart card to smart card. Other companies have developed smart card readers for computers. With such readers in place, electronic money could move online. See generally Karen Kaplan, *E-commerce may help Americans learn to love 'smart' cards*, L.A. TIMES, Oct. 11, 1998, at C1.

29. See Macintosh, *supra* note 14, at 738-39.

30. See *id.* at 750.

31. See *id.*

32. See *id.* at 756-57.

33. See *id.* at 758. Some may question the utility of common units of account, arguing that computer software could be developed to translate any national unit of account into any other national unit of account. However, the transaction costs involved might be higher than expected; the software would have to be constantly updated as exchange rates for over two hundred national currencies fluctuated.

rencies will serve as stable *stores of value*. Competition will drive unstable products out of the market.³⁴ Unlike national monies, private currencies will not be subject to the inflationary monetary policies of national governments and the special interests they represent.³⁵

Some might argue that we could achieve the same advantages with government monies. For example, if the Federal Reserve Board issued its own electronic money, dollars could become the currency for the entire Internet.³⁶

However, the Federal Reserve Board does not plan to issue electronic money at this time.³⁷ More importantly, the Framework reminds us that Internet commerce should be facilitated on a *global* basis. A global marketplace should not depend on the currencies of sovereign nations and the politics and inflationary monetary policies that come along with them.³⁸ If allowed to lead, the private sector can develop stable electronic currencies

More importantly, practical experience teaches that common units of account provide informational benefits to users. Prior to introduction of the euro, the price charged for the same goods varied widely between European countries. *See Faster Forward*, THE ECONOMIST, Nov. 28, 1998, at 83. An investment bank studied fifty-three homogeneous goods across Europe and found that on average, prices differed from the mean by twenty-four percent—a variation twice as large as in the United States. *See id.* This variation was due in part to price opacity. Even when exchange rates were steady, comparing prices stated in differing national currencies was not easy for consumers. *See id.*

On January 1, 1999, Europe entered a new era. Two hundred and ninety million people in eleven countries adopted a common currency, known as the euro. *See Anu Mahmud, Birth of Euro: Impact on Economy*, HONG KONG STANDARD, Jan. 27, 1999. The new unit of account introduced price transparency, allowing buyers to comparison-shop freely across Europe. As experts had predicted, *see Faster Forward, supra*, the introduction of the euro revealed that some prices were out of line, and brought prices down. *See Charles Fleming, The Euro's Arrival Leads One Firm to Cut its Prices*, WALL ST. J. EUR., Jan. 28, 1999, at 13.

34. *See Macintosh, supra* note 14, at 762.

35. *See id.* at 763-64.

36. *See Joshua B. Konvisser, Note, Coins, Notes, and Bits: The Case for Legal Tender on the Internet*, 10 HARV. J.L. & TECH. 321 (1997).

37. Apparently, the Fed fears that direct competition between the government and private sector could stifle the current environment of experimentation and innovation. *See Hearing Before the House Comm. on the Judiciary*, 105th Cong. 619 (June 3, 1998) (statement of Laurence H. Meyer, Member, Board of Governors of the Federal Reserve System).

38. National governments can print large amounts of money to cover deficits, reduce unemployment, or to redistribute income and wealth between creditors and debtors. *See Lewis D. Solomon, Local Currency: A Legal & Policy Analysis*, 5 KAN. J.L. & PUB. POL'Y 59, 66 (1996). Of course, such action results in inflation. *See id.*

that are free of political entanglements and offer the benefits of universal media of exchange, global units of account, and stores of value.³⁹

IV. HOW CAN WE FREE UP THE SYSTEM?

The last of my three questions is the most important for policymakers. What, if anything, can government do to encourage private companies to develop global electronic currencies, or other electronic payment systems for the Internet?

In November 1998, the U.S. Government Working Group on Electronic Commerce issued its First Annual Report on progress made in achieving the goals stated in the Framework.⁴⁰ In the pages that follow, I offer some constructive advice for the Working Group in three areas: proposed federal regulations, uniform law projects, and outdated federal and state laws.

A. Federal regulations

The Framework calls upon regulators to refrain from imposing “inflexible and highly prescriptive regulations and rules” that could inhibit the development of new systems for electronic payment.⁴¹

To a remarkable extent, the federal government has heeded this call. The Federal Reserve Board has declined to extend Regulation E⁴² to electronic stored value products.⁴³ As a result, the development of consumer protection policy for electronic stored-value products has been left to the marketplace. Similarly, the Consumer Electronic Payments Task Force has declined to recommend specific regulations for electronic payment systems, recommending instead that market participants develop policies and

39. See Macintosh, *supra* note 14, at 761-64.

40. See *supra* note 11.

41. See FRAMEWORK, *supra* note 1, at Part I.2.

42. The Electronic Fund Transfer Act, 15 U.S.C. §§ 1693-1693r (1998), regulates electronic fund transfers involving consumers. Regulation E implements the Act. See 12 C.F.R. § 205 (1999).

43. See BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM, REPORT TO THE CONGRESS ON THE APPLICATION OF THE ELECTRONIC FUND TRANSFER ACT TO ELECTRONIC STORED-VALUE PRODUCTS (1997), available at <http://www.bog.frb.fed.us/boarddocs/RptCongress/efta_rpt.pdf>. This report does not recommend any specific course of action. However, it concedes that benefits to consumers might not outweigh the costs of applying Regulation E to electronic stored value products. See *id.* at 75.

procedures to address areas of consumer concern.⁴⁴ According to the Task Force, government should limit its role to providing consumer financial education, monitoring industry developments, and encouraging industry to self-regulate.⁴⁵

Unfortunately, however, there is one federal agency that has resisted the call to laissez-faire. Pursuant to the Money Laundering Suppression Act of 1994 ("MLSA"), Congress amended the Bank Secrecy Act ("BSA") to require any business engaging in money transmitting services to register with the Financial Enforcement Network ("FinCEN") of the U.S. Department of the Treasury.⁴⁶ To implement this mandate, on May 21, 1997, FinCEN issued proposed amendments to the BSA regulations.⁴⁷ Under the amendments, the term "financial institution" would include "money services business,"⁴⁸ which, in turn, would include issuers and sellers of stored value⁴⁹ and money transmitters.⁵⁰ If adopted, the amendments would eliminate any lingering doubt that those who offer or operate advanced electronic payments systems are subject to the BSA.⁵¹

44. See CONSUMER ELECTRONIC PAYMENTS TASK FORCE, REPORT OF THE CONSUMER ELECTRONIC PAYMENTS TASK FORCE 59 (April 1998). In reaching this conclusion, the Task Force voiced its concern that comprehensive new regulation of electronic money could quash competition and innovation, retard the development of a promising new industry, and increase the cost of new products unnecessarily. See *id.* at 59.

45. See *id.* at 59-60.

46. See 31 U.S.C. § 5330(a)(1) (1998); Lee S. Adams & David J. Martz, *Survey: Developments in Stored-Value Cards and Cyberbanking*, 53 BUS. LAW. 1085, 1091 (1998).

47. See Proposed Amendment to the Bank Secrecy Act Regulations, Definition and Registration of Money Services Businesses, 62 Fed. Reg. 27,890 (1997) (to be codified at 31 C.F.R. § 103).

48. See *id.* at 27,897 (to be codified at 31 C.F.R. § 103.11(n)(3)).

49. "Stored value" would include funds or monetary value represented in digital electronics format and stored or capable of storage on electronic media in such a way as to be retrievable and transferable electronically. See *id.* at 27,898 (to be codified at 31 C.F.R. § 103.11(vv)). FinCEN intends this broad term to encompass not only stored value cards, but "other advanced payment system products." See *id.* at 27,893; Linda Noonan, *Many New Businesses May Become Subject to the BSE*, 7 NO. 11 MONEY LAUNDERING L. REP. 1 (1997).

50. See Proposed Amendment to the Bank Secrecy Act Regulations, Definition and Registration of Money Services Businesses, *supra* note 47, at 27,897-98 (to be codified at 31 C.F.R. § 103.11(uu)(3)-(5)).

51. See *id.* at 27,893.

Compliance with the BSA and its regulations is burdensome and expensive.⁵² By increasing cost and effort, the proposed amendments could slow—or even stop—the development of global electronic currencies, and other innovative electronic payment products.⁵³

Moreover, some regulations and programs could embroil electronic payment systems in political controversy. For example, FinCEN plans to exempt transactions involving stored value and other advanced electronic payment products from suspicious transaction reporting⁵⁴—but not for long. Already, the agency has invited comments about the manner in which suspicious transaction reporting should apply to transactions involving stored value products.⁵⁵

Critics of the proposed regulations have questioned whether stored value is within the scope of the MLSA and its grant of authority to FinCEN.⁵⁶ Congress did not discuss stored value when the MLSA was under consideration.⁵⁷ Nor was any evidence produced at that time to demonstrate that stored value or similar electronic payment products were being used to launder money.⁵⁸

52. See Adams & Martz, *supra* note 46, at 1091. For example, financial institutions subject to the BSA must report currency transactions over \$10,000 and keep records on funds transfers over \$3,000. See *id.*

53. See *id.* at 1092; *Uniform Non-Depository Providers of Financial Services Act: Hearing Before the National Conference of Commissioners on Uniform State Laws Drafting Committee* (Oct. 24, 1997) (testimony of Mark E. Plotkin, Partner, Covington & Burling, on Behalf of Mondex USA), available at <<http://www.law.upenn.edu/library/ulc/ndpfsa/plotkin.htm>> [hereinafter Plotkin Testimony].

54. On May 21, 1997, FinCEN published its Proposed Amendment to the Bank Secrecy Act Regulations, Requirement of Money Transmitters and Money Order and Traveler's Check Issuers, Sellers, and Redeemers to Report Suspicious Transactions, 62 Fed. Reg. 27,900 (to be codified at 31 C.F.R. § 103). These proposed amendments would extend to "money services businesses" a suspicious transaction reporting regime that is similar to the one imposed on banks, thrifts, and credit institutions. See *id.* at 27,900-01. However, the amendments would exempt transactions that involve only the issuance or facilitation of transfer of stored value, or the issuance, sale, or redemption of stored value. See *id.* at 27,908 (to be codified at 31 C.F.R. 103.20(a)(4)).

55. See *id.* at 27,904.

56. See, e.g., Plotkin Testimony, *supra* note 53.

57. See *id.*

58. See *id.* Years later, no case of "cyberlaundering" has been detected. See FINANCIAL ACTION TASK FORCE ON MONEY LAUNDERING, 1997-1998 REPORT ON MONEY LAUNDERING TYPOLOGIES, pt. II(ii), *New Payment Technologies*, at 7 (Feb. 12, 1998), available at <<http://www.ustreas.gov/fincen/typo97en.html>>.

Three years ago, when the Federal Reserve Board first proposed applying Regulation E to electronic stored-value products,⁵⁹ Congress required the Fed to study and report on whether Regulation E would adversely impact the cost, development, and operation of such products.⁶⁰ Similarly, Congress should direct FinCEN to conduct a more extensive study to determine whether regulation under the BSA could have a harmful impact on the cost, development, and operation of electronic payment systems.⁶¹ Like the Fed, FinCEN should be asked to consider whether allowing competitive market forces to shape the development of electronic payment systems would more efficiently achieve the objectives of the BSA.⁶² If the answer is “yes,” then the proposed amendments should not be adopted.

B. Uniform law projects

Federal regulators are not the only source of “inflexible and highly prescriptive regulations and rules”⁶³ that could inhibit the development of global electronic currencies. Consider, for example, the Uniform Money Services Business Act (“UMSBA”).⁶⁴ Designed to combat money laundering, the UMSBA would subject money services businesses to a complex system of licensing, examination, reporting, and civil and criminal penalties. “Money services business” includes a person who sells, issues, or provides payment instruments,⁶⁵ including stored value instruments.⁶⁶

59. See 61 Fed. Reg. 19,696 (May 2, 1996).

60. See Economic Growth and Regulatory Paperwork Reduction Act of 1996, Pub. L. No. 104-208, § 2601, 110 Stat. at 3009-469.

61. See Noonan, *supra* note 49, at 4-5.

62. See Economic Growth and Regulatory Paperwork Reduction Act of 1996, *supra* note 60. For example, Mondex USA has pledged to impose low limits on the amount of value that may be stored on consumer smart cards. Merchant smart cards will be rendered incapable of transmitting value to anyone other than as a legitimate consumer refund or as a deposit in a bank. Moreover, Mondex will monitor transaction activity in its system, searching for abnormal patterns of behavior. See Plotkin Testimony, *supra* note 53.

63. See generally FRAMEWORK, *supra* note 1, at Part I.2.

64. National Conference of Commissioners on Uniform State Laws, Uniform Money Services Business Act (March 1999), available at <<http://www.law.upenn.edu/library/ulc/moneysrv/msb399.htm>>. In prior drafts, the UMSBA was known as the Uniform Nondepository Providers of Financial Services Act. The name change was adopted on the ground that “money services business” better described the entities regulated under the Act, and was consistent with FinCEN terminology. See *id.* § 101, Reporter’s Note.

65. See *id.* § 102(17).

66. See *id.* § 102(20). The current draft includes the following definition of stored value instrument:

The term also includes a “money transmitter” who engages in the business of receiving money for transmission or transmitting money.⁶⁷

The drafters have included stored value products within the scope of the UMSBA on the reasoning that the use of stored value as a means of payment is similar to money transmission as a process.⁶⁸ The drafters also are considering whether electronic currency that is transmitted over the Internet falls within the current definition of money transmitter, or needs to be separately addressed in the Act.⁶⁹

The drafters have good intentions but are on the wrong track. The question is not whether emerging electronic payment systems bear some resemblance to money transmission as a process, but whether those systems are mature enough to sustain the burden of uniform legislation at this time. For two reasons, the answer is “no.” First, we have had little or no practical experience with stored value, let alone electronic currencies or other Internet payment systems.⁷⁰ Under such circumstances, drafting becomes guesswork:

[I]t is virtually impossible to draw sensible statutory definitions as to whom should be required to be licensed under the stored-value provisions of a Uniform Act.... Even such deceptively simple terms as ‘issuer’ and ‘redeemer,’ when applied to stored-value, can mean vastly different things among the dramatically distinct types of stored-value systems struggling to emerge in the marketplace today; in such circumstances, any definitions will be so highly specific to one or another type of provider as to be meaningless.⁷¹

Second, unlike more traditional forms of money transmission, smart cards, electronic currencies, and other innovative payment products are struggling to get off the ground. It may be years before these products are

[A] card or other tangible object for the transmission or payment of money which contains a microprocessor chip, magnetic stripe, or other means for the storage of information, which is prefunded, and for which the value is decremented upon each use, but does not include a card or other tangible object that is redeemable only by the issuer in the issuer’s goods and services.

See id. § 102(26).

67. *See id.* § 102(17).

68. *See id.* § 102(20), Reporter’s Note.

69. *See id.* § 102(18), Reporter’s Note.

70. *See Plotkin Testimony, supra* note 53.

71. *See id.*

firmly established in the marketplace. Imposing burdensome laws during this critical period in time could delay or even preclude the emergence of the electronic payment systems that Internet commerce needs. In the spirit of the Framework and the minimalist approach it advocates, the Working Group should ask the National Conference of Commissioners on Uniform State Laws to remove smart cards and other innovative electronic payment products from the scope of the UMSBA.

C. Outdated federal and state laws

Finally, the Framework states that “[e]xisting laws and regulations that may hinder electronic commerce should be reviewed and revised or eliminated to reflect the needs of the new electronic age.”⁷²

Global electronic commerce requires a high level of innovation. Unfortunately, existing laws often place limits on who can innovate. For example, many states prohibit anyone other than a licensed bank from conducting a banking business.⁷³ “Banking” is then defined so broadly that smart cards and other electronic payment products may be included.⁷⁴ In effect, this precludes non-banks from innovating in the payment systems area.⁷⁵

Worse, some existing laws and regulations seem to prohibit innovation altogether. For example, during the Civil War, coins were scarce. Responding to the crisis, Congress authorized the use of postage stamps as currency.⁷⁶ To secure a monopoly for the stamp currency, Congress added the following provision:

Whoever makes, issues, circulates, or pays out any note, check, memorandum, token, or other obligation for a less sum than \$1, intended to circulate as money or to be received or used in lieu

72. FRAMEWORK, *supra* note 1, at Principles (emphasis added).

73. For example, New York prohibits any corporation other than a national bank, unless expressly authorized by the laws of New York, from issuing notes or other evidences of debt to be loaned or put into circulation as money. See N.Y. Banking Law § 131 (McKinney 1998); see generally Anita Boomstein, *Business or Banking?*, CREDIT CARD MANAGEMENT, Sept. 1998.

74. See *id.*

75. For example, American Express or some other money services business that enjoys widespread market recognition and trust might issue electronic money. The money could be denominated in dollars, or independently. Either way, it would represent a claim against a private company. See *supra* note 19.

76. See Thomas P. Vartanian et al., *Echoes of the Past with Implications for the Future: The Stamp Payments Act of 1862 and Electronic Commerce*, 67 BNA'S BANKING REPORT 464 (1996).

of lawful money of the United States, shall be fined under this title or imprisoned not more than six months, or both.⁷⁷

The Civil War is long over, and the postage stamp currency gone. Unfortunately, the currency monopoly lives on, in the form of this provision. If Internet commerce generates a demand for low-cost information services, buyers will need electronic currencies capable of handling micropayments. But, so long as this antiquated statute stays on the books, few may dare to issue a circulating electronic currency that could be used to make micropayments of less than one dollar.⁷⁸

State laws can be equally problematic. For example, California Penal Code Section 648 provides:

Issuing or Circulating Paper Money. Every person who makes, issues, or puts in circulation any bill, check, ticket, certificate, promissory note, or the paper of any bank, to circulate as money, except as authorized by the laws of the United States, for the first offense, is guilty of a misdemeanor, and for each and every subsequent offense, is guilty of a felony.⁷⁹

This dinosaur was enacted in 1872.⁸⁰ Given the reference to *paper* money, the statute may not apply to electronic payment systems. Still, words like “ticket,” “certificate,” and “promissory note” are ominously vague. Laws of this kind, which threaten to make felons out of innovators, endanger our commercial future.

In sum, the Working Group must push to eliminate laws that could interfere with the free development of electronic payment systems. A useful first step would be to identify every federal and state law that could block innovation. Thereafter, the Working Group should work with Congress and the states to encourage repeal of antiquated and obstructive laws.

77. 18 U.S.C. § 336 (1998) (emphasis added).

78. An argument can be made that “obligation” was never intended to include electronic money, and thus, the statute should not apply. After all, Civil War era lawmakers could not have had electronic payment products in mind when they passed the Stamp Payments Act. However, Congress did amend the Act as recently as 1994. *See* Vartanian, *supra* note 76. Unfortunately, this legislative activity could be taken as a sign that Congress intended to breathe new life into the Act, reaffirming and extending its prohibition to all twentieth century obligations—including electronic ones. *See id.* Because the Stamp Payments Act imposes criminal penalties for violations, uncertainty as to the meaning of the word “obligation” could chill innovation. The Working Group should urge Congress to repeal this provision.

79. Cal. Penal Code § 648 (West 1999).

80. *See id.*

V. CONCLUSION

In order for global electronic commerce to achieve its full potential in the new millennium, the Internet needs more than credit cards. It needs cash equivalents, including global electronic currencies capable of transcending national politics and monetary policies. To encourage innovation along these lines, government must not only resist the temptation to unleash new laws and regulations, but also work to repeal the legislative sins of its past.

CLASH OF THE TITANS: REGULATING THE COMPETITION BETWEEN ESTABLISHED AND EMERGING ELECTRONIC PAYMENT SYSTEMS

By Jane Kaufman Winn[†]

ABSTRACT

This article equates the providers of traditional electronic payment services with the Titans of Greek mythology, and the providers of new electronic payment technologies with the Olympians. Professor Winn concludes, however, that unlike the Titans of Greek mythology, these modern Titans appear to be winning in their battle with the upstart Olympians.

This article describes the fundamental characteristics of payment systems, reviews the applicable law, and describes the new technologies that were, until quite recently, expected to displace older electronic payment systems. Professor Winn finds that consumers and merchants, by and large, are happy with the existing regulatory structure. And, because of the failure of new technologies to gain significant market share yet, regulators have not yet been obliged to revise existing regulations to take account of these new technologies.

TABLE OF CONTENTS

| | | |
|-----|--|-----|
| I. | INTRODUCTION: CLASH OF THE TITANS | 676 |
| II. | CHARACTERISTICS OF PAYMENT SYSTEMS | 678 |
| | A. Liquidity | 678 |
| | B. Finality..... | 679 |
| | C. Transaction risk | 680 |
| | D. Systemic risk..... | 680 |

© 1999 Jane Kaufman Winn.

[†] Associate Professor, Southern Methodist University School of Law, Dallas, Texas. Web site: <<http://www.smu.edu/~jwinn>>. E-mail: <jwinn@mail.smu.edu>. The author gratefully acknowledges the thoughtful comments of Russell B. Stevenson, Jr., Ronald Mann, Mitchell Grooms, Kawika Daguio, Dwight Arthur, and Vadim on earlier drafts of this article.

The title of this article is taken from the film, CLASH OF THE TITANS (MGM-UA 1981). No Titans actually appear in the movie, which deals instead with the myth of Perseus and Andromeda. The movie starred Lawrence Olivier, Harry Hamlin, Burgess Meredith, Ursula Andress, Claire Bloom and Maggie Smith, and featured stop-motion animation effects by Ray Harryhausen.

| | |
|---|-----|
| III. TRADITIONAL PAYMENT SYSTEMS: THE TITANS | 682 |
| A. Check clearing | 682 |
| B. Wholesale funds transfers | 684 |
| C. Credit cards | 686 |
| D. Consumer funds transfers | 688 |
| E. Secure Electronic Transactions ("SET") | 689 |
| IV. PAST FAILURES: THE FIRST OLYMPIAN ASSAULT | 691 |
| A. Early fatalities: First Virtual and DigiCash | 691 |
| B. Walking Wounded: MilliCent, CyberCoin, Mondex | 693 |
| V. THROUGH A GLASS DARKLY: THE NEXT GENERATION OF ELECTRONIC PAYMENT SYSTEMS | 695 |
| A. Secure Sockets Layer ("SSL") | 695 |
| B. E-Check | 697 |
| C. InstaBuy | 698 |
| D. Portals | 699 |
| E. NACHA | 699 |
| F. What the future holds | 700 |
| VI. INTERESTS OF STAKEHOLDERS AND GATEKEEPERS IN ELECTRONIC PAYMENT SYSTEMS | 702 |
| A. Stakeholder Interests | 703 |
| B. Gatekeeper Interests | 704 |
| VII. CONCLUSION | 708 |

Banks are dinosaurs. We can bypass them.

— Bill Gates¹

The reports of my death are greatly exaggerated.

— Mark Twain²

I. INTRODUCTION: CLASH OF THE TITANS

One Greek myth tells the tale of an earlier generation of gods, the Titans, who fought with a later generation of gods, the Olympians, for sovereignty over the earth.³ The Titans were the offspring of Uranus, the sky, and Mother Earth. The Olympian gods were the children of Cronus, the king of the Titans. Mother Earth had helped Cronus overthrow his father and seize power. Cronus feared that one of his own sons would likewise dethrone him, so he swallowed each of his own children as soon as they were born. It was not until the sixth child, Zeus, was born that Cronus's

1. Michael Meyer, *Culture Club*, NEWSWEEK, July 11, 1994, at 38.

2. Cable from Mark Twain to The Associated Press (June 2, 1897).

3. This account is taken from 1 ROBERT GRAVES, *THE GREEK MYTHS* 37-44 (1955).

wife, the Titaness Rhea, conspired with Mother Earth to find a way to save the child from his father. When Zeus was grown, with the help of his mother and his wife, he tricked Cronus into drinking an emetic potion, causing him to vomit up his first five children. They were now fully grown, and these Olympian gods joined Zeus in his battle to wrest control of the earth from Cronus and the Titans. The Olympians and Titans fought bitterly for ten years, but the Olympian gods finally won, as had been prophesied by Mother Earth.

The explosive impact of new communications and information processing technology in recent years has triggered a conflict for market dominance between established and emerging players that resembles the clash of the Titans and the Olympians in sound and fury, if not in cosmic relevance. In the marketplace for electronic payment services, the Titans are played by the regulated financial institutions and the Olympians are played by the emerging payment technology start-ups. The role of Mother Earth is played by consumers, who have not yet made the decision to throw out the Titans and bring in the Olympians. As a result, regulators have no clear mandate. Rather, they face the ambiguous and complex task of maintaining the efficacy of existing regulatory efforts in the face of changing circumstances while not discriminating unfairly between either Titans or Olympians.

In the first round of the struggle for dominance in the market for payment services supporting Internet-based commerce, the Titans have emerged victorious while the Olympians are in full retreat. While only a few years ago, it appeared that systems built to leverage new global electronic networks such as the Internet might sweep away the cumbersome older generation of electronic payment systems, this has not proven to be the case.⁴ Consumers have shown a profound lack of interest in the radically new electronic payment systems developed by the Olympians. U.S. consumers actually seem quite happy with the range of electronic payment options currently available. As long as the Olympians continue to neglect the interests of the consumers and other stakeholders in emerging electronic commerce markets, the less novel, but more serviceable, offerings of the Titans will continue to prevail in the market for electronic payment services.

4. On the side of the visionaries, see, for example, DANIEL C. LYNCH & LESLIE LUNDQUIST, *DIGITAL MONEY: THE NEW ERA OF INTERNET COMMERCE* (1996). On the side of the more jaded observers, see Martin Mayer, *Playing Payments Poker*, *INSTITUTIONAL INVESTOR*, Sept. 1998, at 49.

The Olympians may yet be gearing up for a second assault. This wave, though, may focus on utilizing the existing Titan infrastructure through new, Olympian interfaces. As a countermeasure, the Titans may harness Olympian technology. Regardless of the outcome of this battle, the future seems to be one of hybrid systems, where Titan infrastructure is accessed via Olympian technology.

II. CHARACTERISTICS OF PAYMENT SYSTEMS

Payments are a specialized subset of commercial transactions. Payment systems promote commerce by transferring value quickly and effectively and by imposing a minimum of additional costs or risks on the transacting parties. Payment services support transactions in goods and other services as well as transfers of value between financial institutions and their customers. Payment systems operate at the short-term end of the spectrum for financial services, providing rapid and certain transfers of value. They must be efficient, pervasive, and trustworthy in order to minimize the costs that the payment function adds. Any new electronic payment system technologies must not only offer innovative features, they must continue to meet these basic requirements. A payment system can be described in terms of liquidity, finality, transaction risk and systemic risk.

A. Liquidity

Liquidity is commonly defined as the ease with which an asset can be bought or sold for money.⁵ With the exception of the FedWire,⁶ electronic payment systems are not themselves money, but represent a private substitute for money that is acceptable to the transactors. A private payment system substitute for legal tender has liquidity if other types of assets can be converted into and out of that payment medium without causing significant distortions in the market value of the asset solely attributable to the choice of payment device. This type of payment device liquidity can be achieved when there is a sufficiently large number of transactors in the market and transactors can settle transactions in the payment device without making major modifications in other terms of the transaction in order to do so.

5. See ROBERT A. SCHWARTZ, EQUITY MARKETS: STRUCTURE, TRADING, AND PERFORMANCE 523 (1988).

6. FedWire is the wholesale funds transfer system operated by the Federal Reserve Banks. Since the obligations of the Federal Reserve Banks are the legal tender of the U.S., a credit on the books of the Fed that results from a FedWire transfer is also money in the sense of legal tender in the U.S.

B. Finality

Payment systems differ widely in the degree of finality associated with their use. Final payment is the moment when the payment may no longer be revoked.⁷ The rules governing finality of a particular payment device must be clear and universally applied in order to minimize the transaction costs associated with the choice of the payment device. Certainty about the degree of finality, whether great or small, is an essential element of established payment systems.⁸ Although the importance of finality of payments is obvious when payment transactions are viewed in aggregate as a payment system, rules governing finality may be difficult to enforce in practice because of competing concerns at the level of individual transactions. A payer will normally prefer less finality because it can enjoy the float while the payment is processed, and it can reverse payment in the event of a dispute with the payee.⁹ Merchants obviously prefer a system with more finality, as it gives the merchant the benefit of the float and reduces the risk that settlement will be revoked once made.¹⁰ In addition, the party providing the payment service may wish to make exceptions to finality

7. See LARY LAWRENCE, AN INTRODUCTION TO PAYMENT SYSTEMS 336 (1997). Payment by cash in many respects sets the standard for finality, since once cash has been exchanged for another asset, the purchaser has no power other than persuasion to recover the cash. Many forms of electronic funds transfers have the same high degree of finality, but credit card transactions can be unwound weeks or even months after the transaction has taken place due to the statutory right of the card holder to contest charges that appear on periodic statements. See Truth in Lending (Regulation Z), 12 C.F.R. pt. 226.12 (1998) (special credit card provisions, including liability for unauthorized use), *id.* pt. 226.13 (1998) (billing error resolution provisions). Payment by check is more final than payment by credit card, but nevertheless can still be reversed if the drawer of the check manages to instruct the drawee of the check to refuse payment before the check has actually cleared. See U.C.C. § 4-403 (1996) (customer's right to stop payment).

8. See *Banque Worms v. BankAmerica Int'l*, 570 N.E.2d 189 (N.Y. Ct. App. 1991) (“[The] concern for finality in business transactions has long been a significant policy consideration in this State. In a different but pertinent context, we observed in *Hatch v. Fourth National Bank*, 147 N.Y. 184, 192, 41 N.E. 403 [N.Y. Ct. App. 1895] that ‘to permit in every case of payment of a debt an inquiry as to the source from which the debtor derived the money, and a recovery if shown to have been dishonestly acquired, would disorganize all business operations and entail an amount of risk and uncertainty which no enterprise could bear.’”).

9. See RONALD J. MANN, PAYMENT SYSTEMS AND OTHER FINANCIAL TRANSACTIONS: CASES, MATERIALS AND PROBLEMS 117 (1999).

10. Electronic funds transfers, unlike credit cards, generally have a high degree of finality. For that reason, debit cards as a point-of-sale EFT device are very popular with merchants and much less popular with consumers who are accustomed to enjoying the float when they make payments from their bank accounts using checks. *Id.*

rules for established clients or in other hard cases.¹¹ Thus, unless supported by clear legal rules that leave parties little room for discretion, finality of payment will be difficult to enforce consistently and fairly.¹²

C. Transaction risk

When payment is not in the form of a proffer of legal tender, there is an element of credit risk for the party accepting the payment. Even with a payment of legal tender, there is a risk of error in processing the transaction, or of fraud such as forgery. Transacting parties make choices about which forms of payment are acceptable in part based on these transaction risks. Credit cards may have the lowest form of finality of any major modern form of payment; however, they also have the lowest credit risk for the merchants accepting them, provided the merchant obtains prior authorization for a charge. This is because under credit card system rules, card issuers, not merchants, bear the risk that the cardholder cannot pay.¹³ Paper-based payment systems incorporate processes to reduce fraud losses that historically have kept forgery losses within manageable levels. These include the use of magnetic ink printing, comparison of signatures with specimen signatures, and paper stock variations that are hard to reproduce.¹⁴

D. Systemic risk

What is recognized as money in modern economies is rarely legal tender. It is instead a complex fabric of claims on private organizations that circulate with almost the same degree of acceptance as legal tender. Given the large amount of payments made in reliance on the creditworthiness of private parties, the safety and soundness of participants in the payment system is a paramount concern of regulators. Failure of a major compo-

11. For example, U.C.C. § 4-208(c) provides that a drawee bank may not shift the loss caused by a fraudulent indorsement or alteration of a check onto a third party who negotiated the check if the drawee bank's own customer's negligence made it possible for the fraud to occur. See LAWRENCE, *supra* note 7 at 249.

12. For example, the National Automated Clearing House Association just revised its system rules to provide for the first time for fines for clearing house members who fail to perform their assigned roles within the deadlines established by the rules. See NATIONAL AUTOMATED CLEARING HOUSE ASSOCIATION, 1999 ACH RULES: A COMPLETE GUIDE TO RULES & REGULATIONS GOVERNING THE ACH NETWORK at R2 (1999).

13. For a general explanation of the nature of credit card transactions, see LAWRENCE, *supra* note 7 at 513-515.

14. With advances in desktop publishing, however, many of these traditional obstacles on forgery are eroding rapidly.

ment of the payment system would cause disruption throughout the economy, imposing unforeseen losses on unprepared parties. If the failure is large enough to affect liquidity, it could cause a contraction in the level of economic activity generally. As a result, regulators must evaluate not only the degree to which participants in the payment system observe the system rules but also the likelihood that a system participant would fail to meet its obligations to other system participants. The regulators should also evaluate whether the system could withstand the failure of a major participant.

There is no major payment system in the U.S. today that is not subject to a significant level of supervision.¹⁵ Radically new payment systems may fall outside the purview of existing payment system regulations, so they pose a potential threat to the safety and soundness of existing payment systems if the interface between new and old payment systems is not managed carefully.¹⁶ Furthermore, it is unclear whether regular consumers of regulated financial services are in a position to understand and evaluate the risks of new payment systems, even if existing payment services can be sheltered from the risk of insolvency or illiquidity associated with new payment systems. Because failure of any payment device widely accepted in the market may impose unacceptable levels of costs (even on those market participants who did not accept the payment device directly), the model of self regulation and government restraint advocated in the

15. Payment systems that are part of the banking system, such as the check clearing system and funds transfer systems, are subject to a higher degree of government oversight than payment systems run outside the banking system, such as the credit card system. The credit card industry is still subject to some oversight at the level of card issuer-card holder relations pursuant to the Truth in Lending Act.

16. This represents a variation of "Herstatt risk," which refers to the risk that foreign exchange traders or other money market participants assume when they fulfill their obligations to counterparts without requiring the counterpart to fulfill its obligations simultaneously. Herstatt risk is named after the German bank, Bankhaus I.D. Herstatt K.G.a.A., that was closed in 1974 by regulators before it settled its foreign exchange obligations to other banks. See Raj Bhala, *Self Regulation in Global Electronic Markets Through Rein-vigorated Trade Usages*, 31 IDAHO L. REV. 863, 867 n.7 (1995). If a bank offers its customers a choice of various payments services, and assumes on behalf of its clients the risk that any one of those systems will fail to settle, the bank will be exposed to a risk that obligations under each payment system will not in fact be settled at the end of the trading day due to time zone and operating time differences. This is similar to the risk banks face in foreign currency markets that obligations denominated in different currencies will not be settled at the end of the trading day if there are significant differences between operating times of the payment systems.

Framework for Global Electronic Commerce ("Magaziner Report")¹⁷ must be viewed with some skepticism in the context of payment systems.

III. TRADITIONAL PAYMENT SYSTEMS: THE TITANS

Consumer payment systems in the U.S. are still dominated by traditional paper-based services; only a tiny proportion of total payments are currently electronic. Furthermore, these proportions are expected to change only slowly. In 1997, paper-based payment systems still dominated the market for consumer payment services in the U.S.¹⁸ Wholly electronic payment systems currently account for a trivially small proportion of the total market for consumer payment services.¹⁹

A. Check clearing

Banks in the U.S. have been anticipating the advent of the "checkless society" for decades, hoping to replace expensive check processing services with more efficient electronic funds transfer services.²⁰ Yet checks remain the largest segment of the U.S. payment system, and the volume of checks used continues to increase, although the rate of increase is slowing down. No one anticipates that checks will be displaced by electronic payment systems as the dominant form of payment in the U.S. in the near fu-

17. See, e.g., WILLIAM J. CLINTON AND ALBERT GORE, JR., *A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE* (1997), available at <<http://www.iitf.nist.gov/eleccomm/ecom.htm>>.

18. In 1997, personal checks accounted for 52.37% of the dollar volume of consumer payments, and 32.32% of the total number of transactions, while cash accounted for 17.43% by dollar volume and 40.76% by number of transactions. When other paper-based payment devices such as money orders, travelers checks, official checks (including cashier's checks, teller checks and certified checks) and food stamps are added, 73.36% of all consumer payments by dollar amount, and 75.49% by volume were paper-based. *Consumer Payment Systems*, THE NILSON REPORT, Nov. 1998, at 8.

19. Pre-authorized electronic funds transfers, such as direct deposit of payroll, or payments authorized online or by telephone, accounted for only 2.51% by dollar volume and 1.13% by number of transactions of the total. Card-based payments, which may be electronic or may rely on paper processes, accounted for 24.14% of dollar volume, and 23.39% by transaction volume of the total. Credit cards accounted for the largest amount of card-based payments, with 21.12% by dollar volume and 17.99% of transaction volume, with debit cards at 2.78% and 3.69% respectively. Stored value cards, which in the U.S. consist primarily of phone cards, accounted only for 0.14% of dollar volume and 1.51% of transaction volume, and electronic government benefit cards accounted for only 0.10% of dollar volume and 0.20% of transaction volume. *Id.*

20. See Bill Orr, *The Great Card Question: Will It Be Smart Or Debit?*, ABA BANKING J., Sept. 1998, at 54.

ture. As long as checks are the payment device of choice for both consumers and businesses, regulated financial institutions will have to maintain a considerable investment in the equipment and services required to process huge numbers of pieces of paper.

Businesses and consumers remain committed to checks as a payment device because they are more flexible than electronic payment devices in many respects. Any name can be written on the payee line of a check, including the name of a complete stranger. In principle, checks can be written for any amount, no matter how large or small. The drawer of the check enjoys the float until the check finally clears, and may stop payment until it does.²¹ The bank must normally absorb the cost of forged checks, provided the customer has not contributed to the forgery through his or her negligence.²² Because the volume of data captured during the processing of transactions in the check collection system is minimal,²³ personal data regarding checking account customers and their transactions is less likely to be exploited for secondary uses than the more substantial personal data generated or stored in electronic payment systems such as credit card and debit card systems: Finally, U.S. banks did such a brilliant job of marketing the canceled check as the best record of a transaction that many bank customers will not permit their banks to stop sending canceled checks with their monthly statements. The legacy of this marketing campaign has come back to haunt the banks as they struggle to wean their customers from their dependence on pieces of paper in payment transactions. Before U.S. bank customers will willingly surrender their checking accounts for wholly electronic payment services, they will have to be persuaded that the newer, more efficient electronic services offer an equivalent combination of low price, convenience and insurance against risk of loss.

21. See U.C.C. § 4-403.

22. See U.C.C. §§ 3-401, 3-403, 3-406, 3-418.

23. The only electronic data processed by checks is in the MICR line at the bottom of the check, which reflects the limitations of computers in the 1950s when it was introduced. The MICR line includes the customer's bank's routing number, the customer's account number, the number of the check and the amount of the check. Unlike credit and debit card systems, the MICR line does not permit the payee's identity to be processed automatically. The check collection system permits the automated calculation of the customer's balance, but does not support the retention of the same volume of transaction information by the service provider that the credit and debit card processing systems do.

The development of check guarantee services such as TeleCheck²⁴ have increased the willingness of merchants to accept checks, reducing transaction risks associated with accepting checks and making checks an even more liquid payment device for both consumers and businesses. Systemic risks do not normally arise in the context of the check collection system because of the small average amount of payments made by checks relative to the net worth of the banks providing the service and because of the "midnight deadline" rule, which gives banks more than twenty-four hours to make the decision whether to honor a check²⁵.

B. Wholesale funds transfers

The largest segment of the U.S. payment system, when measured by dollar volume transferred, is the wholesale funds transfer market, which transfers trillions of dollars a day.²⁶ By contrast, the market for electronic funds transfers by individuals is much smaller in dollar volume, though the number of individual check and credit card transactions dwarfs the number of wholesale funds transfers.²⁷ It was in the formation of electronic funds transfer systems that regulated financial institutions first developed the business processes necessary to support the secure electronic transmission of customer instructions regarding funds transfers.²⁸

Although funds transfers conducted over funds transfer facilities maintained by the Federal Reserve Banks were subject to the regulation of the Federal Reserve Board, many funds transfers took place over private systems, such as the Clearing House for Interbank Payment Systems ("CHIPS").²⁹ The entire wholesale funds transfer system was not governed

24. See TeleCheck, *TeleCheck Homepage*, (visited April 20, 1999) <<http://www.telecheck.com>>. See also *For a fee, service takes the worry out of checks*, ARIZ. BUS. GAZETTE, May 11, 1995, at 22.

25. See U.C.C. §§ 4-104(a)(10); 4-214(a); 4-215(a).

26. See 2 FURASH & CO., *BANKING'S ROLE IN TOMORROW'S PAYMENTS SYSTEM*, 46, 61 (1994) (study prepared for the Banking Research Fund on behalf of the Payments System Committee of the Bankers Roundtable).

27. In 1993, there were about 70 million FedWire funds transfers and 61 billion checks processed. *Id.* at 10, 46. There were 7.7 billion electronic funds transfers made in 1993 using ATM and ACH networks. *Id.* at 97. Although 95% of these transfers were made through ATM machines, *id.*, the popularity of direct debit and direct deposit through ACH networks, and point-of-sale EFTs have grown rapidly in recent years.

28. For a discussion of the significance of closed network electronic commerce as it developed in the market for funds transfers, see Jane Kaufman Winn, *Couriers Without Luggage: Negotiable Instruments and Digital Signatures*, 49 S.C. L. REV. 739, 757 (1998).

29. For background on CHIPS, see 2 FURASH & CO., *supra* note 26, at 61-65.

by a clear body of law until U.C.C. Article 4A was promulgated in 1989 and adopted by the states shortly thereafter.³⁰ The Article 4A drafting process resulted in many innovations, even though it drew heavily on the practices that had developed among banks and their customers during the 15 years before the drafting committee was established. While a consensus was not easy to achieve, the community of interests shared by both the banks and their customers permitted the drafting process to find workable compromises on many thorny issues.³¹

Wholesale funds transfers in general, and the FedWire system in particular, are highly liquid payment media with a high degree of payment finality and low transaction risks. Transfers of funds over the FedWire are irrevocable when received and settlement is immediate.³² Systemic risk issues are raised when wholesale funds transfers are accomplished outside the FedWire system, such as through the CHIPS system where net settlement occurs at the end of the trading day. CHIPS maintains a series of controls to minimize the risk that one financial institution may fail to settle its obligations at the end of the trading day, triggering a chain reaction of failures that could cause the collapse of the national financial system.³³ Although U.C.C. Article 4A recognizes the possibility that a funds transfer

30. See Prefatory Note, U.C.C. Article 4A (1995).

31. One basic premise regarding funds transfer law was not modified during the drafting process: the price paid by users of the funds transfer system should not include any substantial risk premium to compensate those harmed by many common forms of fraud or error. Regular users of the wholesale funds transfer system did not want to pay a risk-adjusted price for the service, preferring instead to take responsibility for fraud and error prevention and paying only a nominal price to transfer large amounts of money over a system with a very high degree of finality. This decision to keep risk-pooling at a minimum and force each participant, whether bank or customer, to take primary responsibility for fraud and error losses, is well suited to a wholesale market where only sophisticated players are present. It is a complete departure from the paternalistic risk management model used in the consumer credit card system.

32. See 2 FURASH & CO., *supra* note 26, at 45.

33. These controls include bilateral credit limits, which limit the net amount of credit each financial institution will accept from every other participant, and sender net debt caps, which are limits on the maximum debit position any participant may maintain during a trading day. In addition, the financial condition of members of CHIPS is monitored, and sophisticated technological controls are maintained over communications within the system. See 2 FURASH & CO., *supra* note 26 at 63-65.

system like CHIPS might fail to settle,³⁴ in fact CHIPS has never failed to settle since it began operation in 1972.³⁵

C. Credit cards

Credit cards are more than simple payment devices—they provide direct access to a line of credit. While the credit is important in its own right, it has larger implications as well. The financial institutions that sign up either cardholders or merchants are functioning as financial intermediaries, and by providing that intermediating function, minimize the risk to the transacting parties. The intermediaries charge for these services through a discount rate charged to merchants and through interest charges and service fees charged to cardholders. Because the credit card system is a closed system, in which no one can participate as a cardholder, merchant, card issuer or financial intermediary without agreeing to be bound by the system rules, the fact that a merchant accepts credit cards provides consumers with a guarantee of recourse in the event of a dispute in the underlying transaction. Similarly, the card issuer screens cardholder's creditworthiness for merchants.

The market for credit cards gained momentum in the 1960s when some major banks started marketing them aggressively. The banks that were most aggressive sought to minimize their own risk associated with wholesale distribution of credit cards through the use of overreaching form contract terms. This led to a widespread perception that cardholders needed protections in law. The result was Regulation Z,³⁶ which represents a high water mark in U.S. consumer protection law. Regulation Z prohibits the mailing out of unsolicited credit cards,³⁷ limits cardholder liability for unauthorized transactions that occur before the card issuer is notified of the problem to a flat \$50,³⁸ requires the issuer to send periodic statements to the cardholder and requires the issuer to provide certain dispute and error resolution services.³⁹ The result of these regulations is that credit cards have less finality than almost any other payment device.

34. See U.C.C. Article 4A prefatory note. See also U.C.C. §§ 4A-404, 4A-405 (beneficiary's bank may delay acceptance of a payment order until after it has received payment).

35. See 2 FURASH & CO., *supra* note 26, at 64.

36. 14 C.F.R. pt. 226 (1998).

37. See *id.* pt. 226.12(a).

38. See *id.* pt. 226.12(b).

39. See *id.* pt. 226.13.

The Federal Reserve Board has exercised considerable initiative in maintaining this level of consumer protection in new markets, including mail order and telephone order ("MOTO") transactions.⁴⁰ Under Regulation Z, before a card issuer may charge the cardholder for a transaction, the card issuer is first required to provide a means of identifying the cardholder, such as a signature, photo, or fingerprint on the card itself.⁴¹ In MOTO transactions, which provide the model for Internet retail commerce, the card is not present for inspection by the merchant. Because one of the conditions that the card issuer must establish to charge a cardholder cannot be met, the card issuer may not contest a cardholder's claim that a charge was not authorized.⁴² The card issuer will not be left with the loss, however, because under the system rules of the credit card association, the card issuer will be allowed to charge the transaction back to the merchant who presented it.⁴³

The high level of protections mandated for consumers using the credit card system, as well as the access to credit and the risk intermediation performed by the system, help to explain the current dominance of credit cards as the payment device of choice for retail Internet services. The lack of finality created by the Regulation Z right to contest charges when the cardholder is not satisfied with a purchase⁴⁴ is highly favorable to consumers, as is the right of a consumer to contest a charge as unauthorized.⁴⁵ These rules minimize the transaction risks assumed by consumers in Internet transactions by shifting the risks back onto the merchant, the merchant's bank or the card issuer. The popularity of credit cards among U.S. consumers creates liquidity for merchants accepting credit cards as a form of online payment.⁴⁶ The credit card system provides a form of introduction between two parties who otherwise may have no way to evaluate each other's bona fides. The sophisticated security systems that merchants have

40. See 12 C.F.R. pt. 226, supp. 1, cmt. 12(b)(2)(iii)-3 (1999) (official staff interpretations of Regulation Z).

41. See *id.* cmt. 12(b)(2)(iii)-1.

42. The same analysis applies if the merchant only reads the information recorded in the magnetic strip on the card without examining the card itself. See *id.* cmt. 12(b)(2)(iii)-2.

43. See LAWRENCE, *supra* note 7, at 515-516. If the presenting merchant is not available to take the chargeback, the bank that accepted the charge from the merchant becomes responsible for the chargeback. See *id.*

44. See 12 C.F.R. § 226.12(c).

45. See *id.* § 226.12(b).

46. There are over 1 billion credit cards in circulation in the U.S. See 2 FURASH & CO., *supra* note 26 at 83.

developed during decades of MOTO transactions to keep fraud and error losses to a minimum have been transferred to manage risks with Internet-based commerce. Although credit cards have higher transaction costs stemming from the lack of finality and come bundled with a paternalistic risk-management system, these negatives are apparently more than offset by a larger volume of completed transactions and the significant transaction cost savings to merchants and consumers who do not have to research each other's suitability as a counterparty.

D. Consumer funds transfers

Banks and other regulated financial institutions began to offer electronic funds transfer services to customers through automated teller machines ("ATM") in the 1970s.⁴⁷ At the same time, efforts were made to introduce the use of debit cards at point-of-sale ("POS") payment terminals, although these did not capture any significant market share until the late 1990s.⁴⁸ Following the controversy associated with the aggressive and irresponsible marketing of credit cards and the subsequent consumer protection regulations adopted by the Federal Reserve Board to control those abuses, both consumer protection advocates and financial service providers approached the question of consumer protections for ATM card use with definite agendas. The resulting political confrontation produced the Electronic Fund Transfer Act of 1978⁴⁹ and Regulation E.⁵⁰

These two provisions strike a different compromise between consumer and financial institution interests than Regulation Z. The final outcome was a loss allocation rule that progressively placed more of the risk of loss on the consumer when the consumer does not promptly notify the financial institution after learning that the ATM or debit card had been lost. In marked contrast to the flat \$50 limit on consumer liability under Regulation Z, Regulation E contemplates that a consumer who completely fails to notify a financial institution after losing the card may be exposed to losses without any statutory cap.⁵¹

While banks were able to negotiate a lower level of consumer protection for electronic funds transfers than for credit cards, they later found themselves under considerable public pressure to waive the benefits of that

47. See DONALD I. BAKER AND ROLAND E. BRANDEL, *THE LAW OF ELECTRONIC FUNDS TRANSFER SYSTEMS* ¶1.03 (1996).

48. See *id.* at ¶1.03[5].

49. 15 U.S.C. § 1693 et seq. (1994).

50. 12 C.F.R. pt. 205 (1996)

51. See 12 C.F.R. pt. 205.6.

lower level of protection. In the late 1990s, banks began aggressively marketing debit cards as an alternative to credit cards. These cards, marketed with a credit card brand name, had the same convenience for the consumer as credit cards, but very different consumer protections. Consumers and consumer advocates were shocked by the magnitude of the disparity in consumer protection provisions for debit and credit cards. One fundamental disparity is that while fraud and error losses in credit card transactions take place, initially, at the card issuer's expense, while these losses in debit card transactions take place at the cardholder's expense. The cardholder's account at a financial institution may be depleted and remain depleted while the financial institution makes a decision whether or not to recredit its customer's account. In 1997, Visa USA and MasterCard International announced they would not rely on the full protections accorded debit card issuers under Regulation E, but would voluntarily narrow the gap between the treatment of unauthorized or erroneous transactions for debit and credit cards to bring their business practices more closely in line with consumer expectations. They agreed to limit consumer liability for unauthorized transactions to the same \$50 that applies to credit cards, and to shorten the time a financial institution is permitted to make the decision whether to recredit a customer's account for an allegedly unauthorized withdrawal.⁵²

E. Secure Electronic Transactions ("SET")

Standardization is an essential element of any widely distributed system for electronic commerce, and the development of technical standards for interoperability between vendors of electronic payment services has been no exception. Titans such as Visa and MasterCard have attempted to seize the initiative in innovative electronic payment systems by controlling the development of technical standards. This, in turn, would guarantee a continued role for the Titans' proprietary payment services in the electronic commerce markets of the future. The vast financial resources that the Titans can command have been poured into the development and marketing of the Secure Electronic Transactions ("SET") standard, which, if widely adopted, would permit the use of advanced encryption technology in credit card transactions and would help preserve the dominance of credit cards among electronic payment services.

The promise of asymmetric cryptography, in the form of digital signatures, to resolve some of the new security issues raised by conducting

52. See Jeffrey Green, *When Voluntary Action May Not Be Enough*, CREDIT CARD MANAGEMENT, May 1998.

business over open, public computer networks was widely recognized in the technology community a decade ago.⁵³ Without standards to facilitate the adoption of cryptography, however, there was no way for application developers to realize its promise. Visa and MasterCard initially began work on separate standards for the use of public key cryptography in connection with electronic funds transfers, but in 1996 decided to join forces in developing a single standard for secure electronic transactions, or SET.⁵⁴ In 1997, SET was widely expected to make the Internet safe for electronic commerce by resolving some of the uncertainty that prospective transactors felt about the security of Internet transmissions. SET would require each cardholder, financial intermediary and merchant to use a digital signature certificate in Internet transactions. Consumers would send their signed orders to Internet merchants together with encrypted credit card information that would be passed on, unread, to the financial intermediary. The intermediary would process the credit card authorization and notify the merchant, who would then complete the transaction with the consumer.

The SET system ran into problems almost from the outset, and by early 1999 was no longer in the forefront of discussions about Internet commerce security.⁵⁵ While SET may be a very sophisticated system for improving the security of electronic funds transfers over open networks such as the Internet, it is also a complex and cumbersome one. The specifications for SET were initially driven by technological concerns associated with establishing what would have been the first large-scale public key infrastructure using existing financial networks and the Internet. As a result, the operation of SET procedures seemed likely to be too demanding of existing computers and networks. These problems might occur at the level of the individual consumer's home PC or at the level of the merchants and financial intermediaries who were trying to prepare for large volumes of message flow accompanied by large numbers of complex

53. For a general discussion of the role of cryptography in electronic commerce conducted over open networks, see Jane Kaufman Winn, *Open Systems, Free Markets and the Regulation of Internet Commerce*, 72 TUL. L. REV. 1177 (1998).

54. See SET Secure Electronic Transaction LLC, *About SETCo Secure Electronic Transaction, LLC* (visited Apr. 26, 1999) <<http://www.setco.org/about.html>>. For general information about SET, see SET Secure Electronic Transaction LLC, *SETCo Website* (visited Apr. 21, 1999) <<http://www.setco.org>>.

55. The SETCo website FAQ for merchants and users states that "SETCo expects that many banks will begin offering digital certificates to merchants and cardholders in mid-1998." SET Secure Electronic Transaction LLC, *Frequently Asked Questions* (visited Apr. 26, 1999) <http://www.setco.org/faq_usr.html>. In fact, this did not happen.

computations required for cryptographic functions.⁵⁶ When merchants appeared reluctant to invest in the reengineering required to support such demands on their computer resources and networks, SET was marketed as a solution to the problem of consumer fraud through false claims of unauthorized use of a card as well as to the problem of merchant fraud. SET marketing materials never explained how the use of a digital signature certificate would be analyzed in light of Federal Reserve Board opinions interpreting Regulation Z, however.⁵⁷ Notwithstanding the outpouring of support for SET, by early 1999 it had still failed to garner any significant share of the U.S. market and achieved only a few isolated implementations outside the U.S.⁵⁸

IV. PAST FAILURES: THE FIRST OLYMPIAN ASSAULT

When the Internet was still something of a blank slate upon which one could project all manner of utopian or dystopian notions at will, it seemed obvious that major innovations in payment system technology would be needed to support Internet commerce. None of the early pioneers in this field have yet enjoyed the market success that their radical and innovative approaches would seem to have indicated a few years ago, and quite a few have ceased operations altogether.

A. Early fatalities: First Virtual and DigiCash

In 1994, First Virtual was established to offer secure payments technology to support Internet commerce. In 1996, First Virtual enjoyed a successful initial public offering, positioning itself as an Internet payments company.⁵⁹ Individual subscribers authorized First Virtual to charge their credit cards for an initial allocation of funds to spend. Then subscribers could visit the websites of merchants who have signed up with First Virtual and authorize the merchants to debit their balances by using a PIN

56. SETCo has posted a white paper on its website challenging these claims. See CHRIS LETOQ & STEVE YOUNG, SET COMPARATIVE PERFORMANCE ANALYSIS 25-26 (Nov. 2, 1998) (White Paper prepared by Gartner Consulting for SETCo), available at <<http://www.setco.org/download/setco6.pdf>> (PDF file).

57. The SETCo website FAQ for merchants and users stated "Some of the initial key benefits [to merchants] are ... [r]educed costs associated with fraud." SET Secure Electronic Transaction LLC, *supra* note 55.

58. What may have been the most significant factor undermining SET's prospects for acceptance in the marketplace was the development of the SSL standard for secure message transmission. See discussion *infra*, Part V.A.

59. See Elizabeth Douglass, *IPO Nets a Bundle for First Virtual*, SAN DIEGO UNION-TRIB., December 14, 1996 at C-1

number. The success of alternatives such as SSL eroded demand for the First Virtual service, which never achieved a critical mass of individual or merchant subscribers. In 1998, First Virtual announced the cessation of its payment services and the refocusing of the company on electronic messaging services only.⁶⁰

In the 1980s, David Chaum obtained patents for blind signature cryptographic protocols that might support the development of much more novel and significant alternative payment systems.⁶¹ These protocols permit the creation of electronic tokens that can circulate as money in an online environment without revealing the identity of purchasers using the tokens.⁶² Chaum founded DigiCash in an effort to exploit the commercial potential of this technology as a form of online electronic money, and licensed the system to Mark Twain Bank in the U.S. and to various banks in other countries.⁶³ Individuals wishing to use e-cash to make Internet purchases could draw on their account balances at the licensee bank to download e-coins for safekeeping in a "wallet" on the hard drive of their personal computer. When the individual wished to make a purchase, the software would deduct e-cash from the wallet and transfer it to the merchant.⁶⁴ Although Chaum's system is not premised on the use of a conventional system for clearing and settlement to support it, the Mark Twain application allowed the merchant to transfer the e-cash back to the bank to confirm that it had not been double spent. The bank would cancel the e-coins and credit the merchant's account with the bank.⁶⁵ In November 1998, DigiCash announced that it was seeking Chapter 11 bankruptcy relief in order to reorganize its business activities. Mercantile Bank, Mark Twain Bank's successor organization, had earlier discontinued the service, although some overseas banks were still marketing e-coins.⁶⁶

The promise of the e-coin technology may ultimately be realized in U.S. financial markets, but, consumers are not sufficiently motivated by

60. See Bob Woods, *First Virtual Gives Net Payment System Heave Ho*, NEWSBYTES, July 21, 1998.

61. See BRUCE SCHNEIER, *APPLIED CRYPTOGRAPHY* 115 (2d ed. 1996).

62. See DigiCash, *How eCash Works Inside* (visited Apr. 26, 1999) <http://www.digicash.com/index_e.html>. See generally A. Michael Froomkin, *Flood Control on the Information Ocean: Living With Anonymity, Digital Cash, and Distributed Databases*, 15 J.L. & COM. 395, 460 (1996).

63. See James Gleick, *Dead as a Dollar*, N.Y. TIMES MAG., June 16, 1996, at 26.

64. See *id.*

65. See *id.*

66. See Jeffrey Kutler and Carol Power, *Electronic Commerce: Bankrupt DigiCash to Seek Financing, New Allies*, AM. BANKER, November 10, 1998, at 18.

privacy concerns to create the demand that DigiCash's early promoters expected. As with First Virtual, the total number of subscribers to the e-coin system never achieved a critical mass. Purchasers did not enjoy any float, and instead tied up resources in e-coin balances that did not earn interest and were at significant risk of loss should the security of the wallet on the hard drive be compromised. Merchants also faced a risk of loss due to the threat of double spending, although the clearing system provided by the bank was designed to keep such losses to a minimum. An e-coin system would only be competitive with established payment devices if consumers were willing to accept higher transaction risks and greater finality. Apparently, they were not.

B. Walking Wounded: MilliCent, CyberCoin, Mondex

Various other players in the market for innovative electronic payment services have not ceased operations, although the demand for their products remains unclear. These include "micropayment" technologies such as MilliCent and CyberCoin, and smart card payment devices such as Mondex. Micropayment technologies looked destined for greatness a few years ago based on a common forecast regarding the future of Internet commerce. Because many businesses were reluctant to distribute content free of charge over the Internet and because subscription services are uneconomical for many types of content, it seemed likely that some vendors and purchasers would be interested in small, one-shot transactions at prices below those that current electronic payment systems can easily support. Micropayment technologies, which would permit consumers to download electronic money to a personal computer that could then be spent in small increments with participating merchants, seemed almost certain to catch on.⁶⁷ Millicent⁶⁸ and CyberCoin⁶⁹ were two products developed to exploit this market that have languished in a vacuum of consumer demand. While there may be a market one day for Internet micropayments, it is unclear when that day will arrive.⁷⁰

67. See Kimberly Patch and Eric Smalley, *Drop a Dime Online*, INFO WORLD, Nov. 30, 1998, at 71.

68. See MilliCent Microcommerce System, *Find Out What's New, What's Hot and What's Happening With MilliCent* (visited April 21, 1999) <<http://www.millicent.digital.com>>.

69. See CyberCash, Inc., *The CyberCoin Cash Card User's Guide* (visited April 26, 1999) <<http://www.cybercash.com/mscpc/help/cashcardtoc.html>>.

70. Developers of micropayment technology remain hopeful that it will one day prove to be a "disruptive technology" that challenges successful, established competitors by meeting consumers' unstated or future needs. See CLAYTON CHRISTIANSEN, THE

Mondex smart card technology is another promising technology whose moment of greatness always seems to be at hand yet never quite arrives. The Mondex card is a stored-value card that can take the place of cash by permitting transfers of value onto and off of the card.⁷¹ Unlike the Digi-Cash system as implemented by the Mark Twain Bank, there is no requirement that transactions onto and off of a Mondex smart card be cleared through a central system. This feature reduces the transaction costs of the system while also increasing some of the risks such as forgery or money laundering. Because a smart card does not merely store data but can also perform processing functions, the electronic cash function of Mondex can be combined with other functions to enhance its appeal to consumers and merchants, such as tracking loyalty program credits.

As powerful as the Mondex technology is, this kind of smart card-based electronic payment technology has not achieved any noteworthy successes in the U.S. to date.⁷² Millions of dollars of smart cards were distributed without charge during the 1996 Olympic Games in Atlanta, but utilization rates were very low. In 1998, a smart card pilot on the Upper West Side of Manhattan proved to be a disappointment for the financial institutions sponsoring it. U.S. consumers of electronic financial services are apparently too satisfied with existing alternatives such as checks, credit cards and debit cards to adopt this new technology in large numbers, although smart cards are rapidly gaining popularity in other countries. Compared with smart cards, credit and debit cards offer consumers considerable benefits under existing federal regulations.⁷³

INNOVATOR'S DILEMMA (1997) (discussion of "disruptive technologies"). The sudden popularity of MP3 media files has focused the minds of intellectual property rights owners on the need for finely calibrated metering and payment technologies in Internet commerce. See Whit Andrews, *Microsoft Bets \$15M on Maker of Micropayment Technology*, INTERNET WORLD, March 15, 1999.

71. See Mondex Electronic Cash, *What is Mondex?* (visited April 26, 1999) <<http://www.mondex.com/mondex/cgi-bin/printpage.pl?style=noframescash&fname=../documents/intro1.txt&doctype=genp>>.

72. Smart cards are a popular payment device in Europe, where the technology was developed and in Japan. Checks and credit cards are less popular in Europe and Japan than in the U.S. while payment in cash or devices with the same finality as cash such as debit cards are popular, leading to lower consumer expectations with regard to the finality of payments and their ability to enjoy the float.

73. See discussion of credit and debit card liability rules, *supra* Parts III.C-D.

V. THROUGH A GLASS DARKLY: THE NEXT GENERATION OF ELECTRONIC PAYMENT SYSTEMS

The history of commercial exploitation of the Internet is short.⁷⁴ In light of the successes of radically new concepts such as the browser and the hypertext interface, it once seemed plausible that wholly new payment systems might also catch on like wildfire, completely displacing demand for existing services. As shown, however, existing payment systems are meeting the demand for new electronic payment services with only minor modifications. Novel services are finding it difficult to fulfill the conditions required to make a modern payment system function in any environment.

With major obstacles impeding the Olympians' ability to launch alternative payment systems, successful Olympian strategies may consist of finding ways to harmonize Titan and Olympian approaches to electronic commerce. One possible strategy for getting new payment technologies into the hands of large numbers of users is to create hybrid solutions that combine the best elements of both old and new payment systems technologies.⁷⁵ In fact, a few examples of hybrid services currently exist, some owned by the Olympians, some by the Titans.

A. Secure Sockets Layer ("SSL")

The Secure Sockets Layer ("SSL") standard, developed by the Olympian Netscape as a simple, stop-gap solution pending the development of more sophisticated standards such as SET,⁷⁶ has achieved widespread acceptance as the standard for communicating credit card information over the Internet.⁷⁷ It represents an unusual situation where Olympian technol-

74. Before 1995, the NSF Acceptable Use Policy prohibited commercial use of the Internet. See NAT'L SCI. FOUND., OFF. OF INSPECTOR GEN'L, REVIEW OF NSFNET 79-80 (1993), available at <<http://www.nsf.gov/pubs/stis1993/oig9301/oig9301.txt>> (paginated ASCII document).

75. For example, smart cards might be used as access devices for existing account-based systems such as checking accounts or credit cards. The extra processing functions of the smart card could be used first to enhance the security of older systems in ways that are transparent to end users. Once smart cards are in the hands of end users and large numbers of merchants have installed smart card readers, additional application for this technology could be added, such as tracking merchant loyalty programs or disbursing electronic cash equivalents. See Jeffrey Kutler, *Smart Cards: Mondex Again Trumpets The Need For Open Internet Payment Standards*, AM. BANKER, April 14, 1999, at 13.

76. See discussion *supra* Part III.E.

77. Although SSL began as a proprietary standard, Netscape is taking steps to have SSL recognized as a World Wide Web Consortium and an Internet Engineering Task

ogy has enhanced the position of Titan payment systems in Internet commerce. SSL lacks the elegance and coherence of SET, but it has come to dominate the market for retail electronic commerce by meeting the minimum requirements of the relevant stakeholders.⁷⁸

The SSL standard does not use digital signatures to bind human identities to online communications.⁷⁹ Instead, it relies on the use of a digital certificate to identify a computer, the e-commerce server. The consumer's browser validates the server's certificate, and then uses the public key in the certificate to share a symmetric key with the server. For the remainder of the session, the shared symmetric key is used to encrypt communications between the browser and the server, preventing credit card or other sensitive information from being sent over the Internet in the clear.⁸⁰

The SSL protocol permits payment information to be sent over the Internet, replacing a mail or telephone link between the purchaser and the vendor in the MOTO model, but maintaining an equivalent level of security.⁸¹ The SSL standard has proven highly successful because it does not place excessive demands on the average consumer's home PC and has removed what was one of the most pressing concerns associated with Internet commerce: the public nature of its communications infrastructure.⁸² The SSL standard does not address the security of credit card information once it is on the merchant's e-commerce server, nor does it provide any information about which human being entered the credit card information transmitted to the vendor.⁸³

Purchasers have been happy with the level of protection provided by SSL because the normal consumer-friendly credit card system rules apply. The risk that the merchant may misuse consumer credit card information is not borne by the consumer but by the merchant's bank.⁸⁴ Vendors have

Force standard, which would effectively convert it into an open, public standard. See *whatis*, *What Is ... SSL?* (last modified January 4, 1999) <<http://whatis.com/ssl.htm>>.

78. See Cynthia Morgan, *Dead Set against SET?*, *COMPUTERWORLD*, March 29, 1999 at 74.

79. See *id.* The SSL 3.0 specification does, in fact, permit client certificates to be used, although there are no major retail applications using this feature of SSL today. See Netscape Communications Corp., *SSL 3.0 Specification* (visited Apr. 26, 1999) <<http://home.netscape.com/eng/ssl3/index.html>>.

80. See Netscape Communications Corp., *supra* note 79.

81. See Neal Weinberg, *Digital Dough Fails to Rise*, *NETWORK WORLD*, April 12, 1999 at 47.

82. See *id.*

83. See Netscape Communications Corp., *supra* note 79.

84. See discussion *supra* notes 40-46 and accompanying text.

been happy with the level of protection provided by SSL because the risks of accepting credit card charges over the Internet can be brought in line with the risks of accepting MOTO credit card charges.⁸⁵ A merchant can decide whether to do business over the Internet by estimating both the likely return on Internet business (in light of these higher charges) and the likely total volume of chargebacks.

A cloud on the horizon indicates that SSL may not be a lasting solution for securing Internet credit card transactions.⁸⁶ The volume of disputes associated with Internet purchases seem to be disproportionately large—while Internet transactions account for only 2% of card transactions, by some industry estimates they account for 50% of disputes.⁸⁷ Unless the Federal Reserve Board changes its interpretation of Regulation Z regarding the inability of card issuers and merchants to contest a cardholder's claim that a charge is unauthorized in any transaction in which the card was not available for inspection by the merchant,⁸⁸ merchants will have no choice but to press for improved authentication technology or revisions to Regulation Z.

B. E-Check

The established players have not completely ignored the possibility of using new technologies to gain or maintain market share. The benefits of this approach are that new technologies are harmonized with established business practices. This approach was used with the electronic check, or "e-check", developed by a consortium of major banks.⁸⁹

In 1998, various banks entered into a pilot with the Department of Defense to test this new service and it will be some time before the viability

85. Weinberg, *supra* note 81. For instance, merchant banks and credit card associations have adjusted their pricing structure for accepting credit card charges, with the lowest discount rates applying to material-world transactions where the card is present and can be inspected by the merchant, higher discount rates for MOTO transactions, and the highest discount rates for Internet originated charges. Telephone Interview with Steve Watson, Executive Sales Officer, BA Merchant Services, Dallas, Texas (Apr. 21, 1999).

86. See Rod Newing, *Consumers' Trust Is Growing*, FIN. TIMES (LONDON), March 24, 1999, at 7.

87. E-mail from Lynn Wheeler, technologist with First Data Merchant Services Corp., to author (Mar. 26, 1999) (quoting credit card industry sources) (on file with author). This information should be evaluated in light of the investment Visa and MasterCard have made in SET and has not been independently verified.

88. See, 12 C.F.R. pt. 226, supp. 1, cmt. 12(b)(2)(iii)-3 (1999) (official staff interpretations of Regulation Z).

89. For information about the e-check project see Financial Services Technology Consortium, *echeck*, (visited April 20, 1999) <<http://www.echeck.org/>>.

of this concept for a larger market is determined.⁹⁰ The e-check is an electronic equivalent of a paper check, created through the use of new technical standards and encryption technologies. The e-check standard was carefully developed to mimic the functions of check in an effort to ease burdens to bank customers of switching from paper checks to wholly electronic payment systems.⁹¹ It will remain unclear for some time whether enough financial institutions and businesses that currently transfer funds by check will be willing to invest in the technology required to make the e-check as universally acceptable as the paper check or use e-checks in a sufficiently large number so as to make them as inexpensive to process as paper checks.⁹²

C. InstaBuy

CyberCash has developed a new service, InstaBuy, which is rapidly gaining popularity on retail Internet sites.⁹³ The InstaBuy service permits consumers to enter their credit card information into a secure site maintained by CyberCash and then to authorize the InstaBuy service to release the information to Internet merchants who have also signed up for the service.⁹⁴ InstaBuy eliminates the need for consumers to download wallet software, which is required for the CyberCoin service, or to enter their credit card information each time they wish to make a purchase, which is necessary for most Internet retail sites today. Consumers are not charged for the service, can access their credit card information maintained by InstaBuy from any computer with access to the Internet, and can only use InstaBuy with merchants that have been enrolled in the program. This service is less ambitious than the earlier CyberCoin service as it only offers consumers a simpler and more convenient way to make credit card purchases over the Internet using SSL for a secure connection to the mer-

90. See Financial Services Technology Consortium, *Financial Services Technology Consortium and U.S. Treasury Teaming up on Electronic Check Pilot Program*, (visited Apr. 26, 1999) <<http://www.echeck.org/kitprint/index.html>>.

91. See Financial Services Technology Consortium, *Description of the Electronic Check* (visited Apr. 26, 1999) <<http://www.echeck.org/kitprint/index.html>>.

92. The lack of momentum behind the e-check project may not be due to insufficient demand. The banking industry is currently concentrating most of its effort on surviving the Year 2000 problem. Once banking executives and technologists can return their attention to other issues, interest in e-check may increase dramatically.

93. See First USA Corp., *First USA, Premier E-Retailers Team to Enhance Online Shopping Experience; Internet Merchants Facilitate E-Commerce with VersaPay Digital Wallet* (visited Apr. 24, 1999) <http://www.instabuy.com/press/99mar1fusa_versa.html>.

94. See CyberCash Corp., *Instabuy—How It Works* (visited Apr. 24, 1999) <http://www.instabuy.com/check_it_out.html>.

chant's site. Like any other endeavor, however, InstaBuy's ultimate success will depend on whether enough consumers and merchants can be signed up for the service to justify the cost of the service to merchants.

D. Portals

In 1998, portals attracted a great deal of attention as a way to permit end users to find Internet resources more easily, and to increase traffic to Internet commerce sites.⁹⁵ Online financial services have been a common feature of portals.

One electronic payment system likely to be promoted through portals is Internet bill presentment and payment services. With such a service, a consumer could access all his or her monthly bills, review the contents of the bill and authorize payment of the bills using a variety of payment devices. It is unclear at this time whether the market for Internet bill presentment and payment will be a service offered primarily by banks offering Internet services, or technology firms such as CheckFree.⁹⁶ Such firms provide a unified interface to the consumer, but use a variety of payment devices to execute the consumer's payment orders. The terms under which the partnering of Internet access providers, technology firms and regulated financial institutions will take place are still in flux, so it is unclear how many services will be provided by entities outside the scope of current regulations. The magnitude of changes needed in existing regulations will only become clear when the market for Internet bill presentment and payment is better defined.

E. NACHA

In 1998, in an effort to maintain the relevance of the automated clearing house ("ACH") payment system in a world of Internet-based commerce, the Internet Council of the National Automated Clearing House

95. A portal is "a Web site that offers a broad array of resources and services such as e-mail, discussion forums, search engines, and on-line shopping malls." Vincent James & Erin Jansen, *NetLingo: The Internet Language Dictionary* (visited Apr. 22, 1999) <<http://www.netlingo.com/lookup.cfm?term=portal>>. In this sense, a portal is an enhanced version of an earlier generation of Internet search engine or index sites. An end user might be permitted to designate which Internet resources the user would like to be offered access to from a customized portal site, for example permitting an end user to access a home banking site, an online brokerage site, a news site, an online auction site and various Internet retail sites. Portals should reduce the amount of time the end user is required to spend locating and accessing resources, and distributors of goods and services over the Internet to build relationships with their customers.

96. See CheckFree Corporation, *CheckFree Experience Online*, (visited Apr. 22, 1999) <<http://www.checkfree.com>>.

Association ("NACHA") conducted a pilot to determine the compatibility of the ACH system with a browser-based interface for retail Internet commerce.⁹⁷ One of the workhorses of electronic payment services in the U.S., the ACH network performs functions such as direct deposit of payroll and automatic funds transfers for routine expenses such as installment loans or utilities. To date, though, it has not played a major role in the development of Internet-based commerce. The pilot program involved a variety of banks, technology vendors, and members of the NACHA Internet Council in simulated transactions. The pilot demonstrated that while it might be possible to use digital signatures and a browser interface to initiate funds transfers through the ACH system to support retail commerce, further study was needed before such a service could be offered to the public.⁹⁸

F. What the future holds

The market for electronic payment services remains crowded with competing vendors and putative standard-setters, none of which have yet gained a commanding lead over the pack of aspirants. Titans may ensure their continued success in electronic payment services markets if they can persuade Olympians to join them in collaborative relationships. It is possible that Olympians will survive and prosper by licensing their innovative technologies to Titans who have the capital and established relationships with consumers needed to achieve critical market share for new technologies. Such collaborations will permit regulators to maintain current levels of safety and reliability for new products by reviewing the new technologies and their proposed uses in light of existing regulatory standards.

One opening for collaboration between vendors of new technologies and established providers may come if the existing electronic payment system security infrastructure is upgraded or replaced. Electronic funds transfer systems such as those used for retail ATM machines rely on the Data Encryption Standard ("DES") established in the 1970s.⁹⁹ Groups that oppose the current U.S. policy of limiting the use of encryption in the private sector to older standards such as DES have focused on breaking DES

97. See The Internet Council, *The Internet Council Home Page* (visited Apr. 21, 1999) <<http://internetcouncil.nacha.org>>.

98. See The Internet Council, *Authentication & Network of Trust Work Group (ANT) Work Group Description*, (visited Apr. 27, 1999) <<http://internetcouncil.nacha.org/wgdesc.htm>>.

99. DES was developed by IBM as the Lucifer algorithm in the early 1970s, and was adopted as a federal standard in 1976, and as a private sector standard for financial institutions in 1981. See SCHNEIER, *supra* note 60, at 266-7.

as a way of demonstrating the inadequacy of current U.S. government policies regarding encryption.¹⁰⁰ Financial institutions are in the process of replacing DES-based technology with higher levels of security such as double-DES, triple-DES and other encryption standards. At the same time that encryption standards are being reevaluated, it is possible that other elements of the existing payment system infrastructure may also be upgraded. Such replacements may be beneficial—for example, replacing magnetic stripe cards with smart cards would permit the use of more advanced authentication technology in electronic payment systems.

Looking further ahead, services such as financial electronic data interchange and electronic bill presentment may finally permit electronic funds transfers to replace checks for routine business payments. This has yet to occur in large part due to the inability to include transaction information with payment information in transactions among all institutions that accept electronic funds transfers. In order for businesses to have the option to include transaction information with any funds transfer, many businesses and financial institutions will have to upgrade their existing funds transfer processes, which they have been unwilling to do. The federal government is setting a faster tempo for changes with its EFT99 initiative.¹⁰¹ This initiative will force many recipients of payments from the federal government and financial institutions to be able to accept funds transfers with transaction information in the very near future.¹⁰²

New entrants in the market for electronic payment services may enjoy greater success outside mature markets such as the U.S. than inside them. In Europe, for instance, consumers are less accustomed to using payment devices that include an element of float for consumers.¹⁰³ As a result, ac-

100. See Electronic Frontier Foundation, *RSA Code-Breaking Contest Again Won by Distributed.Net and Electronic Frontier Foundation (EFF)* (Jan. 19, 1999) <http://www.eff.org/pub/Privacy/Crypto_misc/DESCracker/HTML/19990119_deschallenge3.html>. For a general discussion of U.S. encryption policy and its critics, see STEWART A. BAKER AND PAUL R. HURST, *THE LIMITS OF TRUST* (1998).

101. EFT99 is an initiative of the U.S. Department of the Treasury designed to all Federal payments, except tax refunds, from paper-based payment systems to electronic funds transfers by January 2, 1999. See Financial Management Service, *Electronic Funds Transfer Home Page* (last updated April 14, 1999) <<http://www.fms.treas.gov/eft/index.html>>.

102. See DEPARTMENT OF THE TREASURY, OFF. OF PUBLIC AFFAIRS, *TREASURY FINALIZES ELECTRONIC PAYMENT RULE AS ENROLLMENT NUMBERS CONTINUE TO RISE* (September 25, 1998) available at <<http://www.treas.gov/press/releases/pr2710.htm>>.

103. U.S. consumers are very familiar with such devices. See *supra* Part III.A (check clearing); Part III.C (credit cards).

ceptance of smart cards as a payment device has been much greater. The dismal failure to date in the U.S. market of smart cards as a payment device demonstrates that U.S. consumers are showing a greater degree of sophistication with regard to product features and a greater resistance to change than most Olympians expected, especially in light of the success of innovative new offerings in other areas of Internet electronic commerce.¹⁰⁴ The greatest successes for such new payment devices may ultimately come in markets in developing countries such as China, where there are virtually no alternative electronic payment technologies.¹⁰⁵ In such markets, there may be no business case for rolling out older models of electronic payment systems where the basic infrastructure is still lacking, and consumers may accept the most up-to-date technology available quite happily.

Emerging payment technologies that gain substantial market share in Europe, Japan or developing countries may be able to leverage that market share to reenter the U.S. market on more favorable terms in the future. As U.S. financial services markets become more integrated into global markets, it is unlikely that the flow of standards and products will always be from the U.S. outward. In the near term, however, established services such as credit cards, debit cards and ATM machines that were perfected in the U.S. and other developed countries, may continue to crowd out more technologically-sophisticated alternatives.

VI. INTERESTS OF STAKEHOLDERS AND GATEKEEPERS IN ELECTRONIC PAYMENT SYSTEMS

These larger trends in the market for electronic payment services and in the development of technology will be shaped by the degree to which the various stakeholders—the consumers, merchants, Titans and Olympians—can protect their interests. In many respects these interests are incompatible. The regulators—the gatekeepers to the market—have the

104. The phenomenal successes of many new Internet electronic commerce ventures raise difficult questions for investors and financiers struggling to determine whether these new models really make any sense or are just a reflection of a bubble economy. *See, e.g.*, George Anders, *Nothing ventured ... "You expect big losses and want \$10 million? Sure, we'll consider it,"* WALL ST. J., March 17, 1999, at A1 (discussing the ability of Internet entrepreneurs to obtain millions of dollars of funding for businesses that are unprofitable today and have no prospect of profitability in the near future).

105. *See* Patricia Lamiell, *China to establish smart card project*, AUSTIN AMERICAN-STATESMAN, August 22, 1998, at D5.

ability to moderate the outcomes that competitive forces alone would produce.

A. Stakeholder Interests

By looking at the costs and benefits imposed on customers by various major payment devices, it is easy to understand why customers of regulated financial institutions still prefer checks to electronic equivalents. The benefits customers enjoy as a result of choosing checks are directly offset by the burdens they impose on merchants. Yet until payment service providers and merchants find a way to meet not just their own needs but also the genuine needs of their customers with electronic payment devices, customers will continue to resist wholesale migration to new forms of payment. While a great deal of consumer resistance may be due to poorly designed interfaces or flawed business models for new payment devices, it may also be due in part to a bias on the part of consumers in favor of payment systems with known, manageable risks. Consumers have demonstrated repeatedly that no matter how favorable a new electronic payment system is for merchants and service providers, they will refrain from using the new system unless they perceive it to be as beneficial to them as existing systems.

Internet merchants are enjoying considerable success today with minor modifications to established systems, but if disputes associated with online credit card transactions continue to grow, merchants may become more willing to pay for more sophisticated solutions such as SET. As businesses develop more integrated, comprehensive information technology systems, they will be looking for payment services that can be integrated seamlessly into those systems. In order for Internet merchants to be able to process payment information automatically, a great deal of standardization will need to take place in the way data describing the transaction is attached to the payment information. In addition, businesses would like the payment transaction information to be available for other uses, such as differentiating between valued and undesirable customers, building closer relationships with more profitable customers, and exploiting the commercial value in transaction data via other processes that cannot now be performed by the merchant.¹⁰⁶

106. It is beyond the scope of this paper to discuss the privacy issues raised by these data practices. Many of these uses are not currently regulated under US law today, although it is possible that some businesses will try to compete by offering to exceed legal minimums in their data privacy policies.

Regulated electronic payment service providers need to find new ways to leverage their existing investment in legacy systems, while keeping up with the rapid pace of innovation characteristic of Internet electronic commerce in general. Banks are at risk of finding their market limited to only the most commodified, low profit financial services, while customers build relationships with newer service providers who are able to capture the more differentiated and profitable front end of the payment system interface. Existing electronic payment systems such as the automated clearing house system for electronic funds transfers have complex, cumbersome interfaces but operate with a high degree of reliability and security.¹⁰⁷ If Yahoo!, Microsoft or Amazon.com can design and retain control of a user interface that integrates automated clearing house payment functions with a customized point of entry to a complete range of Internet services, regulated financial intermediaries will not be able to reap any significant profits from the limited range of services they continue to provide. Banks want to avoid seeing their contribution to the next generation of electronic payment systems reduced to some minimal clearing and settlement service. Collaborating with new technology services may be an easy way for traditional payment service providers to give their systems a fast facelift, but such collaborations also carry the risk that the traditional service may lose its current close connection to the consumer customer or retail merchant.

Technology firms seeking to participate in the construction of new payment systems infrastructure will benefit most if they are able to gain market dominance with proprietary standards, and if they are able to establish a right to payment based on transaction volume or value. If technology firms design systems around open, public standards, they will face a much more competitive market for their services. Just as collaboration between traditional payment service providers and technology firms is perilous for the traditional providers, it is perilous for the technology firm if its partners can eventually master the new technology and internally generate the necessary infrastructure instead of outsourcing its production.

B. Gatekeeper Interests

Although wildcat banking was once a legitimate occupation in the U.S.,¹⁰⁸ regulation of financial services at both the state and federal level

107. See, e.g., 2 FURASH & CO., *supra* note 26, at 29-37.

108. When President Andrew Jackson vetoed the rechartering of the Second National Bank in 1832, the individual states became solely responsible for controlling the banking industry in America. During the next thirty

is now inescapable. The will to maintain a comprehensive framework of regulation over financial markets was strengthened in the wake of the banking crisis of the 1930s, and the savings and loan crisis of the 1980s. The regulatory legacy of these crises is one of highly intrusive oversight and major restrictions on the scope of operations.

The common wisdom now generally equates such heavy handed regulation with obstacles to effective competition, and makes the regulatory calculus for government agencies considering intervention in emerging markets complex and ambiguous. At least in the case of credit cards and Regulation Z protections for consumers, however, the heavy hand of regulation may have given established payment systems the competitive edge they needed to achieve rapid dominance of the market for retail Internet payment services. If generations of intrusive government regulation have dulled the wits of consumers to the point they cannot distinguish between acceptable and unacceptable levels of risk associated with new electronic payment services, then regulators will face a difficult task in protecting consumers while not stifling competition. If years of intrusive government regulation have instead produced a generation of consumers who have a reasonably good understanding of the costs and benefits associated with existing electronic payment systems and have a marked preference for those regulated systems which treat consumers well, then market demand may push providers of electronic financial services into the arms of regulators. Before the Titans can profit from the latter scenario, however, many of their existing payment services will need something of a facelift to meet the technological demands of emerging electronic commerce markets.

The new payment service providers may enjoy a distinct competitive advantage over regulated financial intermediaries if they are able to compete head-on without being subject to the regulatory burdens themselves. If the Olympian developers are able to establish significant market share in emerging electronic payment technologies before regulators decide to subject them to traditional financial market regulations, they will enjoy a significant competitive advantage over their regulated Titan competitors. If regulators want to avoid encouraging unregulated competitors from

years, commonly referred to as the era of 'free banking,' state-created currency removed the public's confidence in banks and resulted in the most chaotic era in our nation's financial history."

Stephen G. Stroup, Comment, *Smiley v. Citibank (South Dakota)*, N.A.: *Charging Toward Deregulation of the Credit Card Industry*, 22 DEL. J. CORP. L. 601, 603 (1997) (footnotes omitted).

taking market share from regulated financial intermediaries, they will need to maintain their current agnostic position with regard to regulating innovative offerings by established electronic payment service providers.

Should regulated financial intermediaries choose not to collaborate with technology innovators, but try to compete directly with them for consumer markets, the regulated intermediaries are sure to lobby regulators to release them from some of their current legal and regulatory obligations. If regulators hold financial institutions to high standards that impose substantial costs and limit their ability to keep up with more nimble competitors but fail to maintain an equivalent level of control over those new competitors, regulated financial institutions would experience a loss of market share without any offsetting increase in the effective level of consumer protection.¹⁰⁹

In an era of deregulation, regulators may face considerable obstacles to expanding existing regulatory regimes to cover new payments technologies. In 1996, the Federal Reserve Board ("FRB") issued proposed amendments to Regulation E to deal with new developments such as stored value cards.¹¹⁰ Congress responded by prohibiting the FRB from taking action to finalize any amendments to Regulation E until the FRB had determined whether Regulation E could be applied to stored value cards "without adversely affecting the cost, development and operation of such products."¹¹¹ The FRB, duly chastised, delivered its report to Congress on April 2, 1997, finding that policy statements or education programs might be less costly and just as effective as regulations in protecting consumers interests.¹¹²

Until the technology innovators can mount a more credible threat to the ruling hegemony of highly-regulated electronic payment services, the wait-and-see and incremental-reform approaches currently being taken by regulators are adequate, because the magnitude of the threat posed by emerging services is negligible. If this trend continues in the future, then

109. See, e.g., letter from John J. Byrne, American Bankers Association, to FDIC, FRB, OCC and OTS (Mar. 8, 1999), available at <http://www.aba.com/aba/static/KYC_cmtltr.html>.

110. See Electronic Funds Transfers, 61 Fed. Reg. 19,696 (to be codified at 12 C.F.R. pt. 205) (proposed May 2, 1996).

111. Economic Growth and Regulatory Paperwork Reduction Act of 1996, Pub. L. No. 104-208, 110 Stat. 3009, 3009-469 (1996), § 2601(a)(1).

112. See BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM, REPORT TO THE CONGRESS ON THE APPLICATION OF THE ELECTRONIC FUNDS TRANSFER ACT TO ELECTRONIC STORED-VALUE PRODUCTS (1996), available at <http://www.bog.frb.fed.us/boarddocs/RptCongress/efta_rpt.pdf>.

regulators in this field may be spared the task of finding new techniques for domesticating rapid technological innovation without stifling competition. Competitive offerings from regulated providers of electronic payment services may permit consumers to continue to rely on payment services that manage a number of the risks of electronic commerce fairly and efficiently as well as advancing the technology of market infrastructures.

Instead of resorting to regulation, the problem of competing proprietary solutions for meeting the future needs of consumers in Internet commerce might be resolved through the work of standards-setting organizations.¹¹³ For example, the current network of retail electronic funds transfers conducted through ATMs and point-of-sale terminals using debit cards operates today due to the standardization work of the American National Standards Institute ("ANSI") Accredited Standards Committee ("ASC") X9.¹¹⁴ The work of open standards-setting organizations might be helpful in resolving some of the current uncertainty regarding the future direction of electronic payments technologies. However, formal standards-setting organizations such as ANSI ASC X9 are competing with a wide range of private and more informal standards-setting organizations that hope to influence the future of electronic payments technologies. For example, the Financial Services Technology Consortium¹¹⁵ and the Banker's Roundtable Banking Industry Technology Secretariat¹¹⁶ are financial services industry trade associations that cater to large banks and financial services companies that have tried to establish themselves as industry leaders through projects such as e-check¹¹⁷ and the Bank Internet Payment System.¹¹⁸ In 1999, however, not many proposed standards were making headway in the market, either because they had not yet moved beyond the pilot stage or because, like SET, they were not being adopted following

113. For a general discussion of standards-developing organizations, see CARL F. CARGILL, *OPEN SYSTEMS STANDARDIZATION: A BUSINESS APPROACH* (1997).

114. For information about the work of the ANSI ASC X9, see Accredited Standards Comm. X9 & American Bankers Ass'n., *X9 Overview* (visited Apr. 26, 1999) <<http://www.x9.org>>.

115. Membership in FSTC requires payment of dues ranging from \$5,000 per year to \$16,000 per year. See Financial Services Technology Consortium, *Join the FSTC Online* (visited Apr. 26, 1999) <<http://www.fstc.org/membership.html>>.

116. Membership in BITS is limited to the 125 largest banks and thrifts in the U.S. See The Banker's Roundtable, *The Banker's Roundtable* (visited Apr. 26, 1999) <<http://www.bankersround.org>>.

117. See discussion *supra*, Part V.B.

118. Financial Services Technology Consortium, *The Bank Internet Payment System (BIPS): Leading the Way to Electronic Commerce* (visited Apr. 26, 1999) <<http://www.fstc.org/projects/bips/>>.

pilot projects. The very large number of other putative standards-setting organizations competing within the electronic payment services market has reduced the probability that any one organization's standard would achieve universal acceptance and recognition, and so enjoy the leverage that the network effects of such standardization would provide.¹¹⁹

VII. CONCLUSION

Two years after the publication of the Magaziner Report, its recommendation that regulators refrain from regulating until a need has been established appears to have been sound even when applied to electronic payment systems. The more utopian projections for the future of electronic payment systems have not yet been realized, and the more dire threats to regulated payment systems have not yet materialized. In 1999, the market is dominated by traditional financial intermediaries offering conventional electronic payment services augmented with minor innovations to adapt to the Internet.

In the original myth of the clash of the Titans, the Titans first came to power with the assistance of Mother Earth in their struggle to overthrow Uranus. The Titan Cronus's efforts to forestall his eventual overthrow were doomed to fail in the face of Mother Earth's support for the Olympians, and resulted only in postponement of the inevitable. In the battle for dominance of the electronic payment systems market, the Titans are clearly still in command and, at least in the U.S., the developers of alternative technologies have yet to establish much of a beachhead. This is far from the final outcome, however, as innovative payment technologies may become established outside the U.S. first, and then capitalize on their position in global markets to reenter the U.S. market on more favorable terms in the future. In addition, innovators may succeed in hollowing out the economic value of the basic payment system franchise, if the services provided by regulated financial intermediaries become more constrained and commodified, and competitors gain control of novel interfaces or delivery mechanisms that narrow the services the payment systems Titans offer.

119. For a list of various electronic payment standard setting initiatives, see European Union Open Information Interchange, *Electronic Payment Mechanisms* (visited Apr. 26, 1999) <<http://www2.echo.lu/oii/en/payment.html>>. For a collection of web resources relating to electronic payment issues, see American Bar Association, *ABA Electronic Financial Services Subcommittee Home Page* (visited Apr. 26, 1999) <<http://www.abanet.org/buslaw/efss/home.html>>.

Consumers have shown a high degree of rationality in their choice of electronic payment systems, and have stayed away from more risky or less favorable innovations. Regulated electronic payment systems offer incidental attributes such as float, or reversibility in the event of dispute. Consumers may migrate toward regulated systems because they provide these incidental benefits without regard to how well systemic risk issues are managed. But so long as regulators guarantee the provision of both, then consumers can migrate toward the most favorable package of rights and obligations and the system will enjoy the oversight necessary to keep systemic risk to manageable levels.

Consumer resistance to radical innovation is hard to interpret, but it is possible that this resistance is based in part on some understanding of the benefits consumers currently enjoy under existing systems and are unwilling to surrender without some equivalent benefit. New systems that provide a big boost to the electronic payment service provider, a marginal benefit to the merchant and negligible benefits to the consumer are failing, while existing electronic payment systems that preserve existing benefits are prospering with only small modifications.

If regulated financial intermediaries continue to meet the demand for electronic payment services through incremental innovations in established services developed under the scrutiny of regulators, then emerging payment systems do not pose as much of a threat to existing payment systems or the economy generally as once feared. Regulators will have the time they need to gauge the real risk posed by innovation to existing systems and to adapt existing regulations to carry forward an equivalent level of protection in new systems.

OLD AND NEW ISSUES IN THE TAXATION OF ELECTRONIC COMMERCE

By David L. Forst[†]

To date, the response of the world's taxing authorities to the rise of electronic commerce can best be characterized as "hurry up and wait." Conferences have been convened; white papers have been issued; commissions have been appointed, but there is still precious little concrete advice that tax advisors can give to their clients.

While black letter rules have yet to materialize, statements of principle have abounded. Certain of these stated principles are more or less self-evident, such as the idea that electronic commerce taxing regimes need to be fair, consistent, administrable, and the like. However, other principles are neither self-evident nor harmonious, suggesting that the real work of developing an electronic commerce taxing regime will be neither easy nor quick.

The principal difficulty in developing an electronic commerce taxing regime is that the Internet is still a new medium whose full ramifications are not close to being understood. Accordingly, at least for now, governments at all levels are not eager to commit to rules that could potentially erode their tax bases. Business is not pushing for new rules either, since existing rules can be interpreted favorably.¹ Thus, it seems that electronic commerce tax law will not be a radical departure from existing rules, but instead will develop piecemeal and reactively to cure perceived taxpayer abuses or revenue misallocations.

This paper will discuss the development of electronic commerce tax law from the perspectives of international taxation and U.S. state taxation—the two most fertile areas. It will identify the firm rules that have dribbled out so far, the principal points of tension that have prevented further

© 1999 David L. Forst.

[†] Attorney, Fenwick & West LLP, Palo Alto, CA.

1. Under existing rules, the right of a jurisdiction to tax usually does not arise unless and until a taxpayer has some sort of physical connection with that jurisdiction. Since the consummation of an electronic transaction does not necessarily require a physical connection with the purchaser or its jurisdiction, businesses can reasonably take the position that they should not be subject to tax under taxing regimes that require a physical presence. See David L. Forst, *The Continuing Vitality of Source Based Taxation in the Electronic Age*, 15 TAX NOTES INT'L 1455, 1467-71 (1997).

progress, and the areas where progress is most likely in the foreseeable future.

I. INTERNATIONAL TAXATION

A. The Continuing Vitality of Source Based Taxation

In the international arena the most concrete principle that has emerged so far is that traditional source based taxation rules continue to be robust. Since the beginning of this century, the world's nations have basically agreed that the country with the primary right to tax the profits of an enterprise is the country in which the enterprise earns the income (the "source country").² In so doing, they have rejected an alternative residence based taxing regime, in which the enterprise's country of residence has primary taxing jurisdiction over business profits. This consensus is reflected most completely in the international network of some 1,500 bilateral income tax treaties for the avoidance of double taxation.³ The issue of residence based taxation eclipsing source based taxation in the context of electronic commerce was brought to the forefront by the U.S. Treasury Department more than two years ago. In November 1996, the U.S. Treasury Department issued a paper entitled *Selected Tax Policy Implications of Global Electronic Commerce*,⁴ which stated that source based taxation could be rendered obsolete by electronic commerce:

The growth of new communications technologies and electronic commerce will likely require that the principle of residence based taxation assume even greater importance. In the world of cyberspace, it is often difficult, if not impossible, to apply traditional source concepts to link an item of income with a specific geographical location. Therefore source based taxation could lose its rationale and be rendered obsolete by electronic commerce.⁵

However, the ascendancy of residence based taxation now appears to be a dead letter. Since the Treasury Department issued its November 1996

2. See The Organization for Economic Cooperation and Development Model Tax Convention on Income and Capital, art. 7(1) (1992), 1 Tax Treaties (CCH) ¶ 191 [*hereinafter* *OECD Model Tax Convention*].

3. See *id.*

4. Department of the Treasury, Office of Tax Policy, *Selected Tax Policy Implications of Global Electronic Commerce* (visited Nov. 1, 1997) <<http://www.fedworld.gov/pub/tel/internet.txt>>.

5. *Id.* at 19.

Report, major international organizations, such as the Organization for Economic Cooperation and Development (“OECD”), have not endorsed residence based taxation over source based taxation, and to this author’s knowledge neither has the taxing authority of any individual government.⁶ Even the Treasury Department itself has not pushed the issue further, and instead (as discussed more fully below) has begun to apply existing tax principles to transactions effected via electronic commerce.

In hindsight, the rejection of residence based taxation was inevitable. Residence based taxation is simply not necessary as a technical matter because traditional source based rules can be applied to transactions effected via electronic commerce.⁷ Furthermore, countries historically have been unwilling to cede their right to tax the foreigner.⁸ This unwillingness is exacerbated in the case of electronic commerce where most of the world’s electronic merchants are, and at least for the foreseeable future will be, U.S. residents.⁹

B. Future Significance of Source Based Taxation

With source based taxation secure, the issue has shifted away from whether taxation at the source should be retracted to whether taxation at the source should be expanded. The world’s taxing authorities could end up concluding that source based principles will need to be liberalized to insure that all countries receive a fair share of electronic commerce tax revenue.

This issue is best illustrated by the tension between two policy goals set forth by the OECD. In October 1998 the OECD’s Committee on Fiscal Affairs issued a report entitled *Electronic Commerce: Taxation Framework Conditions*,¹⁰ which set forth certain broad, but potentially incompatible, policy goals.

One policy goal stated in the OECD Ministerial Report is that existing tax principles should apply to the taxation of electronic commerce:

6. See Organization for Economic Cooperation and Development, *Electronic Commerce: Taxation Framework Conditions* (visited Jan. 8, 1999) <http://www.oecd.org/daf/fa/e_com/frameworkke.pdf> [hereinafter *OECD Ministerial Report*].

7. See Forst, *supra* note 1

8. See *supra*, note 2 and accompanying text.

9. See U.S. Could Gain Income if Current Policies Apply to Electronic Commerce, *Attorneys Say*, TRANSFER PRICING REP., July 16, 1997, at 178.

10. See *OECD Ministerial Report*, *supra* note 6.

The taxation principles which guide governments in relation to conventional commerce should also guide them in relation to electronic commerce. The [Committee on Fiscal Affairs] believes that at this state of development in the technological and commercial environment, existing taxation rules can implement these principles.¹¹

To this end, the OECD has stated that it would clarify the Commentary to its Model Tax Convention to take into account certain issues related to electronic commerce, such as how the current definition of permanent establishment¹² applies where electronic commerce transactions are conducted through a website on a server located in a given country and how income earned from certain electronic transactions should be classified for the purposes of taxation.¹³ Similarly, the U.S. Treasury Department has announced that it will apply traditional title passage rules in determining the source of income from the electronic transmission of inventory property.¹⁴

However, another of the OECD Ministerial Report's policy goals, the maintenance of sovereignty and fairness, could conflict with the goal of preserving existing principles. The Ministerial Report states:

Any arrangements for the application of [existing] principles to electronic commerce adopted domestically and any adaptation of existing international tax principles should be structured to maintain the fiscal sovereignty of countries, to achieve a fair sharing of the tax base from electronic commerce and to avoid double taxation and unintentional nontaxation.¹⁵

What the OECD appears to be saying is that the fiscal sovereignty of countries and a fair sharing of the tax base from electronic commerce is more important than a strict (or possibly even loose) adherence to existing

11. *Id.* at 4.

12. The permanent establishment concept generally provides that a country cannot tax a non-resident unless the non-resident has a permanent presence in the country, e.g., through an office, factory, workshop or the like. The permanent establishment concept was developed before the era of electronic commerce, and thus, as presently conceived, does not take into account whether a country can tax a non-resident if the non-resident's sole contact with the country is an electronic one. See Forst, *supra* note 1, at 1467-71.

13. See Organization for Economic Cooperation and Development, *Electronic Commerce: A Discussion Paper on Taxation Issues* (visited Jan. 8, 1999) <http://www.oecd.org/daf/fa/e_com/discusse.pdf> [hereinafter *OECD Discussion Paper*].

14. See T.D. 8785, 63 F.R. 52971-52983 (1998).

15. *OECD Ministerial Report*, *supra* note 6, at 4.

technical tax principles. As a political matter, the OECD's statement cannot be disputed. Any international tax regime governing electronic commerce will have to be consistently applied, and consistent application will not occur if certain countries believe that they are being denied their fair share of tax revenue.

An issue of particular concern in this regard is disintermediation, or the removal of the middleman from business transactions. An international example of disintermediation is a U.K. person purchasing a book from an Amazon.com website or an airline ticket from a Travelocity website rather than from a local bookseller or a local travel agent. In both cases, the U.K. middleman is cut out, and the U.K. loses tax revenue. However, whether disintermediation will result in a material reallocation of the world's tax revenue is unclear. It has been suggested the process of reintermediation, or the rising of new types of middlemen who help facilitate the online flow of goods and services, will make up for any dislocations caused by disintermediation.¹⁶ But right now, it is simply too early to tell what effect disintermediation or other consequences of electronic commerce will have on worldwide taxation.

As a result, caution is the order of the day. The OECD is unwilling to propose any far-reaching changes to the international tax regime. The OECD's Committee on Fiscal Affairs recently concluded that "it would be premature to reach any conclusion as to the effect of electronic commerce on the sharing of tax revenues between source and residence countries or to put forward alternatives to the present rules of tax conventions concerning the tax of business profits."¹⁷

Industry has also discouraged new rule making. Existing rules prohibit a country from asserting taxing jurisdiction over a foreigner unless that foreigner has some sort of physical presence in the country.¹⁸ Thus, a business with an electronic, but not a physical, presence in a country can reasonably take the position that it should not be subject to taxation in that

16. Online escrow agents are good examples of reintermediators. An online escrow agent facilitates online sales by holding the buyer's cash until the buyer receives and is satisfied with the goods that it has purchased. See Jamie Beckett, *Rise of the Online Middleman; Escrow Services in Demand as Net Auction Sites Proliferate*, S.F. CHRON., Jan. 14, 1999, at B1.

17. See *OECD Discussion Paper*, *supra* note 13, at 25.

18. See, e.g., OECD Model Tax Convention, *supra* note 2 (providing that the profits of an enterprise of a contracting state shall be taxable only in that state unless the enterprise carries on business in the other contracting state through a permanent establishment situated therein).

country. Industry is opposed to new rules because it believes (probably correctly) that any changes to existing rules would lower the threshold on countries' ability to tax the foreigner. IBM Chairman Louis Gerstner recently told the OECD, "I would ... encourage you to consider an official time-out.... Let's commit to allow this new kind of commerce to chart its natural pattern and then explore how a predictable and neutral tax regime might be superimposed on it."¹⁹

The caution expressed by government and industry is compounded by their mutual desire not to retard the growth of electronic commerce. There has been widespread rejection of taxes targeted specifically at electronic commerce. The leading example of this type of tax is a bit tax, which would tax every item of digital data—or "bit"—that flows over the Internet. The OECD has not endorsed this type of tax, and in a recent report the U.S. government trumpeted its role in defeating worldwide endorsement of the tax, stating that it would have discriminated against electronic commerce.²⁰

With the number and variety of goals expressed, change will almost certainly be incremental. The most realistic prospect for change is relaxation of barriers to countries taxing at the source, through, for example, expanded use of consumption taxes or liberalization of the permanent establishment principle in bilateral income tax treaties. This type of change would use existing principles as a starting point and would ameliorate any deviations from the current allocation of the world's taxing revenues. However, this type of change also could subject businesses without a permanent establishment (as that term is currently understood) in a particular country to taxation in that country. Thus, prospects for even this type of change are far from certain.

II. STATE AND LOCAL TAXATION

In the state and local arena even incremental change has been temporarily halted. In October 1998, Congress passed the Internet Tax Freedom Act,²¹ which prohibits states or political subdivisions from imposing from October 1, 1998 through October 21, 2001 taxes on Internet access and

19. Peter Menyasz, *OECD Joint Declaration Stresses Use of Consumption Point as Locus for Taxation*, DAILY TAX REP., 196 DTR G-2 (Oct. 9, 1998).

20. See U.S. GOVERNMENT WORKING GROUP ON ELECTRONIC COMMERCE, FIRST ANNUAL REPORT (Nov. 30, 1998), reprinted in BNA DAILY TAX REP., Dec. 1, 1998, available at <<http://www.doc.gov/e-commerce/E-comm.pdf>>.

21. Pub. L. No. 105-277, 112 Stat. 2681 (1998).

multiple or discriminatory taxes on electronic commerce. The Act also calls for the establishment of a nineteen member Advisory Commission on Electronic Commerce to conduct a study and submit findings in the year 2000. A grandfather clause protects taxes imposed and enforced before October 1, 1998. Accordingly, states and localities are essentially barred from imposing any new types of taxes on electronic commerce until late 2001. In the meantime, there has been no shortage of discussion on the issue.

Like international taxation, a predominant issue in state and local taxation is whether barriers to taxing out-of-state sellers should be relaxed. The current standard, set forth by the Supreme Court, is that a seller must have some physical presence in a state, whether through property, employees or independent contractors, before a state can subject the seller to its taxing laws.²² For example, if a resident of California purchases a book from a local bookstore, the seller has an obligation to collect California sales tax on the transaction. But if the same person purchases the same book from an electronic book merchant who does not have a physical presence in California, the seller does not have an obligation to collect California sales tax on the transaction.

As of now, it is unclear whether, or to what extent, this discrepancy will continue. The current system has been criticized as economically distorting because it provides a tax incentive to purchase goods over the Internet rather than through conventional means. It also has been criticized as inadequate to meet states' and localities' revenue needs. On the other hand, it has been argued that if current rules were relaxed, it would be too burdensome for electronic merchants to comply with the different sales tax regimes of the 50 states and countless localities. Technical issues also would arise, such as where delivery would take place when a seller transmits a good electronically.

In the meantime, states seem to be of the dual view that taxation should not hinder the development of electronic commerce and that electronic commerce should not erode their tax bases. The actions of California in this regard are instructive. In August 1998, California adopted its

22. See *Quill Corp. v. North Dakota*, 504 U.S. 298 (1992). In *Quill*, the taxpayer sold goods to customers in North Dakota through a mail order catalogue. The taxpayer did not have any employees in North Dakota, or otherwise have a physical presence in North Dakota. At issue was the whether the corporation was required to collect North Dakota use tax on sales made to residents of North Dakota. The Supreme Court held that a corporation must have some physical presence in the state for the state to be able to impose the obligation to collect use taxes.

own Internet Tax Freedom Act,²³ which imposes a moratorium on certain taxes relating to electronic commerce. The legislature stated that its intent was to place the greatest possible barrier on the creation of discriminatory taxes.²⁴ Consistent with this sentiment, the California Franchise Tax Board ruled that a person having a mere electronic presence in a state, e.g. through a server, does not have sufficient nexus with the state to be subject to the state's taxing laws.²⁵ However, the Electronic Commerce Advisory Council, appointed by former California governor Pete Wilson, recommended in December 1998 that the federal government compel out-of-state vendors to collect sales tax under a simplified, streamlined system that would reduce compliance burdens.²⁶

Thus, the outlook for state and local taxation seems to be some sort of relaxation in the rules restricting the taxation of remote sellers, but nothing is certain in this regard. It is certain that nothing material will be accomplished until at least 2001 after the current federal moratorium is lifted.

III. CONCLUSION

With black letter rules on electronic commerce taxation in short supply, the onus is squarely on governments at all levels to develop a coherent and well thought-out taxing regime. A prerequisite for the development of international rules is the agreement of the world's nations on what those rules should be. Specifically, the world's nations need to agree on how the basic principles of international taxation—that income should be subject neither to double taxation nor nontaxation—should be adapted to income derived from international electronic transactions. Such a consensus will not develop if certain nations believe that they are being denied their fair share of electronic commerce taxing revenue. Thus, it is likely that a worldwide consensus will not develop at least until the world's nations are secure in their understanding of how the Internet will effect their economies.

23. CAL. REV. & TAX. § 65001 (West 1998).

24. See Kathleen Wright, *States of Mind: Logging onto the New Tax World: California Passes its Tax Freedom Act*, 98 STATE TAX NOTES 197-3 (1998).

25. See, e.g., CAL. CODE REGS. tit. 18, § 1684 (West 1999), which provides that the use of a server on the Internet to create or maintain a Web page or site by an out-of-state retailer is not considered as a factor in determining whether the retailer has substantial nexus with California.

26. See Jeremy Holmes, *Council Issues Report on E-Commerce, Urging Simpler Rules, Out-of-State Collection*, 240 DAILY TAX REP. H-1 (1998).

Under present state and local sales and use tax law electronic merchants who have no physical presence in a particular state need not collect sales and use tax on behalf of that state. Federal action will be necessary for states and localities to collect more vigorously sales and use tax. Such action will not be forthcoming until at least late 2001 when the current federal moratorium on the imposition of Internet-related taxes expires. In the meantime, businesses will continue to interpret existing rules to their advantage, and state and local governments will continue to leave valuable sources of revenue untapped.

THE SPEED GAP: BROADBAND INFRASTRUCTURE AND ELECTRONIC COMMERCE

By Howard A. Shelanski[†]

ABSTRACT

Although high-speed, broadband telecommunications services are not yet widespread outside of urban and commercial areas, they are starting to reach an increasing range of residential customers. Greater availability of high-speed communications links is likely to increase the growth of electronic commerce and other Internet applications, to the benefit of consumers and online businesses alike. Regulation of advanced services may, however, affect the speed of residential broadband deployment and the prices for such services in the short run. This essay discusses some important legal constraints underlying current regulatory proceedings and the impact those constraints may have on the spread of affordable broadband services.

TABLE OF CONTENTS

| | | |
|------|--|-----|
| I. | AN OVERVIEW OF BROADBAND AVAILABILITY TO CONSUMERS | 722 |
| A. | Current Deployment of Advanced Network Capability..... | 723 |
| B. | Broadband Options in the “Last Mile” | 724 |
| 1. | Telephone Network Solutions: ISDN and DSL..... | 725 |
| 2. | Cable Network Solution: Cable Modems..... | 727 |
| 3. | Wireless and Satellite Solutions | 728 |
| 4. | Where the Residential Market Is—and Where It Needs To Be. | 729 |
| II. | THE IMPORTANCE OF BROADBAND CONNECTIONS FOR E-COMMERCE..... | 731 |
| A. | The Consumers’ Perspective: Lowering Search Costs | 731 |
| B. | The Sellers’ Perspective: Reducing Barriers to Entry..... | 732 |
| C. | The Advantages of Broadband and the Challenge for Telecommunications. | 735 |
| III. | REGULATION OF BROADBAND SERVICES AND IMPLICATIONS FOR E-COMMERCE... .. | 736 |
| A. | Background of the Advanced Services Proceedings..... | 736 |

© 1999 Howard A. Shelanski.

[†] Acting Professor, University of California, Berkeley, Boalt Hall School of Law (on leave 1998-1999); Senior Economist, Council of Economic Advisors. I am grateful to Bert Huang, the editors at the *Berkeley Technology Law Journal*, and to participants in the conference on the Legal and Policy Framework for Global Electronic Commerce, March 5-6, 1999 at the University of California at Berkeley. The views expressed in this essay are the author’s and are not necessarily shared by the Council of Economic Advisors or any other government agency.

| | |
|--|-----|
| B. The FCC's Advanced Services Proceedings..... | 738 |
| C. The Regulatory Outlook: Increased Competition at the Cost of Cheaper Speed?..... | 739 |
| IV. CONCLUSION..... | 744 |

Telecommunications infrastructure is critical to the growth of electronic commerce. Telephone networks, cable systems, and other providers of facilities are essential intermediaries that can shape the volume and nature of transactions between online buyers and sellers. The faster and less expensive the links are between users and the Internet, the more quickly electronic commerce is likely to grow. Competition, innovation and regulatory changes have all contributed to the development of a more efficient, higher capacity telecommunications network that is increasingly well suited to moving large amounts of data quickly. There is, however, a point at which broadband transmission stops: the local, residential network. The extension of broadband capability beyond its current scope to a majority of small businesses and households is an important challenge for the communications industry.

Part I of this essay will discuss the current state of broadband capability in U.S. telecommunications networks. Part II will then discuss the importance for electronic commerce of increasing residential access to advanced, high-speed telecommunications services. Finally, Part III will examine how statutory constraints and tradeoffs underlying current regulatory proposals might affect the availability and affordability of residential broadband services. It suggests that the 1996 Act may constrain the Federal Communications Commission ("FCC") to favor rules that maximize the number of competitors in the broadband market at the expense of rules that maximize the spread of low-priced, advanced service offerings to residential customers.

I. AN OVERVIEW OF BROADBAND AVAILABILITY TO CONSUMERS

This section will begin by discussing changes in the telecommunications system's ability to provide high-capacity lines to customers and to process information in digital format, both of which are essential for broadband services. It will then discuss how, because of the high costs of deploying fiber lines to most individual customers, several technologies have been developed to increase capacity of the communications plant that telephone and cable carriers have already constructed. It will argue that deployment of those technologies—namely integrated services digital network ("ISDN"), digital subscriber line ("DSL"), and cable modem

service—has helped to make broadband service cheaper and more widely available, but not yet on a ubiquitous scale to residential consumers.

A. Current Deployment of Advanced Network Capability

Substantial progress has been made in upgrading telecommunications infrastructure to meet the needs of the information sector of the economy. When AT&T was broken up in 1984, not one “central office”—the offices where the switches that route telephone calls are located—contained advanced, digital signaling technology. By 1997, over 97 percent of central offices deployed such technology,¹ and over 99 percent of customer lines were routed through such switches.² Similarly, in 1984 only a very small number of links used to transport telephone traffic between central offices were made of fiber optic cable; the vast bulk were low-capacity copper lines.³ By 1990, 60 percent of interoffice transmission links were fiber, and by 1997 the proportion of fiber transport plant had reached nearly 96 percent.⁴

FCC figures show that from 1993 through 1997, overall deployment of high-capacity, fiber optic cable in the U.S. telephone system increased from 2.3 million to 3.4 million miles in long-distance networks,⁵ from 6.6 million to 12.2 million miles in incumbent local telephone networks,⁶ and from 0.2 million to 1.8 million miles in competitive local exchange networks.⁷ Total fiber mileage increased an estimated 16 percent in 1997 alone, and actual fiber capacity by the end of 1998 was almost certainly much higher.

While the paving of the “Infobahn” has reached the freeways and main roads, it has not yet reached the neighborhood streets. For the most part, the high-capacity fiber infrastructure stops well short of individual customer lines—often called “loops,” or the “last mile”—that connect individual customers to the network. Of the 150 million customer lines oper-

1. See INDUSTRY ANALYSIS DIV., FEDERAL COMMUNICATIONS COMM’N, TRENDS IN TELEPHONE SERVICE 90 tbl.17.2 (July 1998).

2. See *id.*

3. Indeed, in 1986, total fiber deployment by AT&T was less than 30 percent of its total network, including long distance lines where the bulk of fiber was used. See John Haring & Ewan Kwerel, *Competition Policy in the Post-Equal Access Market*, 62 Rad. Reg. 2d (P & F) 587, n.18 (OPP Working Paper, Feb. 1987).

4. See INDUSTRY ANALYSIS DIV., *supra* note 1, at 91 tbl.17.3.

5. See JONATHAN M. KRAUSHAAR, FEDERAL COMMUNICATIONS COMM’N, FIBER DEPLOYMENT UPDATE END OF YEAR 1997 10 tbl.2 (1998).

6. See *id.* at 24 tbl.6.

7. See *id.* at 36 tbl.14.

ated by the Bell operating companies (the major incumbent carriers), 86 percent were copper and only 14 percent were fiber at the end of 1997.⁸ Because some competitive local exchange carriers have been building all-fiber networks, the percentage of fiber loops for the overall market may be slightly higher than the percentage for the incumbents' networks alone. But the competitive carriers have only about 3 percent of the local market by lines,⁹ so the total percentage of customer lines served by fiber loops is still almost certainly under 20 percent.

Not only is the proportion of fiber loops small, but the distribution of those links is heavily skewed toward businesses and urban customers. Once fiber "backbones" are put in place in dense areas, as they have been in many cities, it can be economical to build a fiber link from the backbone to an office or apartment building. The distances are short—often a matter of yards—and a single building will either have multiple customers or a very high-revenue customer. The economics of building fiber links to customers in less dense areas are much less promising. Loops are much longer—a matter of miles rather than yards—and at the end of that loop generally lies one, relatively low-revenue customer. As a result, no carriers are currently building fiber lines to individual customers outside of the densest urban areas.¹⁰

B. Broadband Options in the "Last Mile"

The absence of fiber deployment to individual customers means that the speed of data transport drops precipitously at the point where information is handed off from the network's transport lines to the customer's loop. Given the time and cost required to build out fiber networks, the solution for bringing broadband service to residential customers must, in the foreseeable future at least, work over existing residential infrastructure: either the copper phone loops or the coaxial links of the cable television network. In addition to solutions based on landline telephone and cable systems, wireless technologies may also become important in the residential broadband market. Today, three technologies that meet the constraints of existing facilities are beginning to enter the market for residential

8. INDUSTRY ANALYSIS DIV., *supra* note 2, at 91 tbl.17.3.

9. See COUNCIL OF ECONOMIC ADVISORS, PROGRESS REPORT: GROWTH AND COMPETITION IN U.S. TELECOMMUNICATIONS 1993-1998, 24 (Feb. 8, 1999) (White Paper).

10. See KRAUSHAAR, *supra* note 5, at 21 n.18.

broadband: ISDN line service and DSL service over the telephone network,¹¹ and cable modem service.

1. *Telephone Network Solutions: ISDN and DSL*

Two ways of providing broadband transmission over copper telephone lines at modest cost are now in use. These technologies, ISDN and DSL, differ in their capabilities and in how they make use of existing infrastructure. ISDN allows transmission rates up to 128 kilobytes per second (kbps) over the circuit-switched voice network, about twice the best rate achievable by conventional modems.¹² Using an ISDN modem is just like using a regular computer modem in that each use requires dial-up to the telephone network. According to FCC data, by 1997 about 40 percent of local telephone company central offices, where the main switches that serve customer lines are located, were capable of providing ISDN service.¹³ Those central offices together serve about 93 million customer lines, roughly 70 percent of the total in the United States.¹⁴ Residential ISDN prices have recently fallen to as low as \$25 per month (not including Internet access), with initial set-up charges of \$125 plus the cost of an ISDN modem.¹⁵ ISDN's drawbacks include potentially high usage payments, frequent difficulty in achieving maximum bit rates, and the lack of an "always-on" connection that can be used without the delay of a dial-up process.¹⁶

DSL service overcomes some of the drawbacks of ISDN because it bypasses the circuit-switched voice network by routing data traffic to a packet-switched network. This allows more economical always-on connections and much faster speeds. By using modems that divide copper phone lines into separate bands for data traffic, DSL achieves download

11. At the high end of the telecommunications market are high-capacity data links called T1 (or T3) lines. Prices vary by distance, contract length, and share of line capacity, with the minimum monthly charge being around \$300. See *Telco Express* (visited Mar. 2, 1999) <<http://digiquote.telcoexpress.com>> (providing an online pricing tool for digital line rates around the country based on location and distance). Because T1 lines tend to be affordable only for large businesses and institutions, they are not considered part of the solution for real consumer-level broadband service—i.e., service affordable by households and small businesses.

12. See *Digital Starter: ISDN*, COMPUTER SHOPPER, Feb. 1999, at 300.

13. See INDUSTRY ANALYSIS DIV., *supra* note 1, at 90 tbl.17.2.

14. See *id.*

15. See *Digital Starter*, *supra* note 12, at 300. Usage charges are 1 to 2 cents per minute in addition to the monthly fee.

16. See Richard Sekar, *A Panacea for DSL Access*, TELEPHONY, Jan. 18, 1999.

speeds from 128 kbps to 7 Mbps.¹⁷ DSL service is not yet widely available, but that is changing. By the middle of 1998, DSL service was available to at least some consumers in about 30 states,¹⁸ and various providers have announced aggressive plans to expand the reach of their DSL offerings. For example, incumbent local exchange companies are pursuing different strategies, but are aiming to serve between 24 and 70 percent of their customers by 2000.¹⁹ In addition, competitive local exchange companies focusing on data services have entered a number of markets. Altogether, independent analysts predict that by 2000, over 40 million U.S. households will have access to DSL service.²⁰

Prices for DSL have started to fall accordingly. Bell Atlantic offers DSL service with Internet access at prices starting as low as \$40 per month, plus an installation charge recently listed at over \$400.²¹ Pacific Bell now offers DSL service, including Internet access, for as low as \$39 per month for 384 kbps speeds; installation and necessary equipment require an additional one-time fee of just under \$200.²²

Although DSL is promising and becoming more widely available, several technical issues limit the number of customers with access to the

17. See Shawn P. McCarthy, *Internet Technologies to Watch*, LOGISTICS MGMT. DISTRIBUTION REP., Jan. 31, 1999, at 74.

18. See Memorandum from Carol W. Wilner, Director, Federal Government Relations, AT&T, to author, (Feb 5, 1999) (on file with *Berkeley Technology Law Journal*) (providing maps depicting states with DSL and cable modem service).

19. SBC, which currently serves 37 million customers, is targeting 8 million for broadband availability by 2000. See *Got Bandwidth? Pacific Bell Answers California's "need for speed" with \$39 ADSL Service, Major Availability* (visited Apr. 9, 1999) <http://www.sbc.com/PB/News/Article.html?query_type=article&query=19990112-04>. Bell Atlantic, which serves 42 million customers, has set approximately the same goal. See *Bell Atlantic.net cuts price of Infospeed DSL package* (visited Apr. 9, 1999) <<http://www.ba.com/nr/1999/Mar/19990331001.html>>. Bell South began service in 1998 with 7 cities, and plans to offer service in 30 cities total by 2000. See *Fastaccess city availability* (visited Apr. 9, 1999) <http://www.bellsouth.net/external/adsl/city_availability.html>. U S West currently provides service in 40 cities. See *U S WEST Brings Lightning Fast New Internet Access to Homes in 40 Cities by June 1998* (visited Apr. 9, 1999) <<http://www.uswest.com/com/insideusw/news/012998.html>>. Ameritech plans to make broadband available to 70 percent of its 21 million subscribers by 2000. See *Ameritech Launches High Speed Internet Service* (visited Apr. 9, 1999) <http://www.ameritech.com/media/release/view/0,1038,8421_2,00.html>.

20. See Wilner, *supra* note 18.

21. See *Bell Atlantic Infospeed DSL Pricing* (visited Mar. 2, 1999) <http://www.bell-atl.com/adsl/more_info/pricing.html>.

22. See *Fastrak DSL—Pricing & Availability* (visited Mar. 2, 1999) <<http://www.pacbell.com/products/business/fastrak/dsl/pricing.html>>.

service: transmission over DSL lines is generally effective only for customers located a short distance, generally within about three miles, from a central switching office.²³ Performance of DSL transmission declines with loop length, but also varies with condition of the loop and quality of equipment attached to the loop; older copper loops that have been patched and repaired over the decades will often have to be reconditioned before they are suitable for DSL transmission.²⁴ Technological advances are starting to provide improvements, but for now DSL remains an option primarily in areas where loops are short and in good condition.

2. *Cable Network Solution: Cable Modems*

The hybrid fiber-coaxial plant of cable systems also has broadband capacity and can be configured for two-way, high-speed data service through the use of cable modems.²⁵ As originally built, however, that pipe runs one-way, toward the consumer; to provide broadband service over the cable network, the plant must be upgraded to two-way capability for the more interactive applications of the Internet or for voice services.²⁶ The investment for such upgrades is substantial, and by one estimate only about 15 percent of systems have been converted.²⁷ But the natural high-speed capacity of cable systems, and the fact that cable is readily available to 98 percent of American households, make it a natural and, for residences, the leading broadband competitor.

Cable systems currently provide high-speed data services to about 300,000 customers, but are expanding aggressively. By the middle of 1998, cable modem service was available to some households in 44 states.²⁸ Since 1995, \$18 billion have been invested in cable upgrades,²⁹ and it is predicted that cable modem service will be available to over 40 million households by 2000.³⁰ AT&T's merger with Tele-Communications, Inc. ("TCI") is premised on upgrading TCI's cable sys-

23. See, e.g., Eric Krapf, *Slow roll for DSL*, BUS. COMM. REV., Aug. 1998, at 47.

24. See *id.*

25. See generally GEORGE ABE, CISCO SYSTEMS, RESIDENTIAL BROADBAND 180-90 (1997) (discussing the principles of operation for cable modems).

26. To be sure, the "upstream" channel away from the consumer need not be as big as the downstream channel, but some upstream capacity is necessary.

27. See James B. Speta, *Handicapping the Race for the Last Mile* 27 (Feb. 16, 1999) (unpublished manuscript, on file with *Berkeley Technology Law Journal*).

28. See Wilner, *supra* note 18.

29. See *id.* (citing Credit Suisse First Boston, *The Infrastructure Report*, Dec. 15, 1998).

30. See *id.* (citing various analyst reports from the Yankee Group).

tems to serve up to 18 million customers with high-speed Internet access within the next few years.³¹ Recently, prices for cable modem service have fallen to about \$40 per month excluding Internet access.³²

3. *Wireless and Satellite Solutions*

Finally, wireless solutions may also be just over the horizon. The wireless services that are likely to provide broadband data capability are not, however, the cellular telephone and personal communications service ("PCS") technologies with which most consumers are familiar. Even with digital conversion of the wireless telephone networks in the U.S. over the past several years, the data rates those systems support are less than the copper, landline network.³³

More promising for broadband purposes are land-based (as opposed to satellite), fixed wireless systems like multichannel multipoint distribution service ("MMDS") and local multipoint distribution service ("LMDS"). These systems use microwave transmission technology to send signals over a 30-70 kilometer radius.³⁴ They have the advantage of low start-up costs³⁵ and by 1997 there were 73 MMDS operators serving 1 million video customers in the United States.³⁶ MMDS and LMDS systems have some drawbacks: they require line-of-sight transmission paths and are subject to interference—even from bad weather.³⁷ MMDS is the more established of the two systems, and is estimated to pass over 30 million homes, although only about 1 million subscribe to MMDS video serv-

31. See *AT&T and TCI complete merger* (visited Apr. 9, 1999) <<http://www.att.com/press/item/0,1193,382,00.html>>.

32. See Scott Bernard Nelson, *Life on the Internet Fast Lane*, *KIPLINGER'S PERS. FIN. MAG.*, Jan. 1999, at 117.

33. See, e.g., Douglas N. Knisley, et al., *cdma2000: A Third Generation Radio Transmission Technology*, *BELL LABS TECH. J.*, July/Sept. 1998, at 65 (noting that the current generation CDMA technology—a radio transmission standard for PCS service—can provide data speed transmission of under 14.4 kb/s at best).

34. See ABE, *supra* note 25, at 343.

35. See *id.* at 347 (noting that in Los Angeles, a single MMDS antenna can reach upward of 4 million households, making the infrastructure investment less than \$20 per residence in the coverage area).

36. See *id.* at 344.

37. See, e.g., *In re Annual Assessment of the Status of Competition in Markets for the Delivery of Video Programming*, FCC CS Docket 98-102, para. 81 (Dec. 23, 1998) (discussing line-of-sight problems) [hereinafter *Annual Assessment of the Status of Competition*]; ABE, *supra* note 25, at 346 (noting MMDS is limited by line-of-sight considerations); Speta, *supra* note 27, at 31 (citing Scott Seidel, *Broadband Wireless Services: In the Line of Sight*, *Bellcore Exchange*, Spring 1997, at 21-22) (noting that rain can affect LMDS service quality at certain spectrums).

ices.³⁸ Its ubiquity is promising, however, and could make MMDS an important broadband entrant if digital compression allows its capacity to increase and if interference and other technical issues can be resolved. LMDS is of much more recent vintage and, although capable of very high-bandwidth transmission, is not considered a near-term entrant into the residential broadband market.³⁹

Finally, satellite services have entered the market and may, as they have in the video market,⁴⁰ prove a powerful competitor for broadband services. Few subscribers to date take advantage of the limited satellite offerings, like DirecPC, now available.⁴¹ But given that satellite broadcasting, or “DBS” service, is moving towards having 15 million subscribers,⁴² and that additional satellite systems have been licensed and are coming online, further offerings are likely in coming years.

Although wireless technologies will likely become more important players in broadband transmission, at present they lag behind other technologies. The most likely near-term solutions to the slow access speeds available to residential customers are those that make use of the landline telephone and cable networks. ISDN, DSL and cable modem services will thus likely see the fastest growth in the near future.

4. *Where the Residential Market Is—and Where It Needs To Be.*

Right now, residential broadband is more promise than reality. Although broadband access is now available in most states, coverage within those states is limited. A recent study found high-speed services offered to selected customers in only 10 percent of counties, although those counties together contain 45 percent of the American population.⁴³ This suggests that advanced services are starting to spread to residents of densely populated areas, but rural customers will have a longer wait. Even if the promises of telephone carriers and cable systems are met, fewer than half of American households will have broadband access in the next couple of

38. See Annual Assessment of the Status of Competition, *supra* note 37, para. 83.

39. See, e.g., Daniel Sweeney, *LMDS: Finally Ready for Prime Time?*, AMERICA'S NETWORK, Aug. 1, 1998, at 22.

40. See, e.g., Howard Shelanski, *Video Competition and the Public Interest Debate*, in TELEPHONY, THE INTERNET, AND THE MEDIA 91, 100 (Jeffrey K. MacKie-Mason & David Waterman eds., 1998).

41. See Les Freed & Frank J. Derfler, Jr., *Hughes Network Systems' Direct PC Internet access via satellite*, PC MAGAZINE, Apr. 20, 1999, at 160.

42. See Annual Assessment of the Status of Competition, *supra* note 37, para. 62.

43. See *State of the Internet: USIC's Report on Use and Threats in 1999* (visited Apr. 16, 1999) <http://www.usic.org/usic_state_of_net99.htm>.

years. But substantial investment is being made in expanding such offerings, and greater availability is inevitable. The two unknown variables are price and speed of deployment.

The best current prices for residential customers are, as indicated above, about \$50 per month for a package of DSL or cable modem service and Internet access. Whether this will be considered affordable by the majority of Internet users is unclear. While the price seems high for those living on the median U.S. family income of about \$30,000, other communications services such as cable television subscription have proven to be fairly insensitive to income.⁴⁴ Broadband access for Internet service might follow a similar pattern, especially if economically bundled with video and voice telephone service. But the economic structure of broadband demand is as yet unknown.

It is very likely, however, that lower prices will substantially increase the spread of broadband subscribership. Indeed, future new purchasers of Internet access may be increasingly cost conscious. The available data indicate that average income of Internet users is declining. In 1995, the average household income of an Internet user was over \$50,000.⁴⁵ The latest Pew Center survey shows that the fastest growing groups of new Internet users are those with much lower income and educational levels than in the past.⁴⁶ The survey finds that 23 percent of new users have annual household incomes below \$30,000 and that 39 percent of new users never attended college.⁴⁷ This is a healthy development, but it also suggests that, over time, customers will be increasingly hard to attract at a given access price. So, for purposes of the growth of e-commerce, the price premium for speed will have to be low enough to reach customers farther down the demand curve for Internet access.

44. See generally Robert Kieschnick & Bruce McCullough, Federal Communications Commission, *Do People Not Subscribe to Cable Television Because They Can Not Afford the Service? A Review of the Evidence* (Aug. 1998) (unpublished manuscript, on file with author).

45. See PROGRESSIVE POL'Y INST., *THE NEW ECONOMY INDEX: UNDERSTANDING AMERICA'S ECONOMIC TRANSFORMATION* 31 (1998).

46. See Pew Research Center for the People and the Press, *Online Newcomers More Middle-Brow, Less Work-Oriented: The Internet Audience Goes Ordinary* (Jan. 14, 1999) <<http://www.people-press.org/tech98sum.htm>>; Bob Tedeschi, *European Union Advances E-Commerce Policies*, N.Y. TIMES ON THE WEB (Apr. 27, 1999) <<http://www.nytimes.com/library/tech/99/04/cyber/articles/26commerce.html>> (citing Yankee Group estimate of \$13 billion).

47. See *id.*

From the perspective of electronic commerce, the challenge for the broadband market is to meet the growth targets announced by carriers, and to do so at prices that not only allow the carriers to make the required return on investment, but also make broadband subscription attractive to a large number of households. As discussed below, the benefits to electronic commerce from such deployment are likely to be substantial for both buyers and sellers.

II. THE IMPORTANCE OF BROADBAND CONNECTIONS FOR E-COMMERCE

The convenience and novelty of online shopping has sparked rapid growth in the volume of electronic commerce. Recent estimates of retail sales over the Internet in the United States range from \$8 billion to \$13 billion for 1998, up from \$3 billion in 1997,⁴⁸ and there seems little reason to believe the market will develop at a slower rate in the near future. Eight million Americans are estimated to have made online purchases this past holiday season.⁴⁹ Established Internet businesses are becoming more user-friendly and sophisticated, while new entrants are coming (and going) at a rapid pace. "Infomediaries" that help consumers to search and sort online businesses have entered the market. And existing infrastructure is, for the moment, supporting substantial growth in the online marketplace. The real question is not whether there will be growth, but what trajectory it will follow. The ability of the telecommunications industry to provide fast and inexpensive pipes between online shoppers and Internet sites is an important factor in the answer.

A. The Consumers' Perspective: Lowering Search Costs

Many factors other than the cost and capacity of telecommunications connections limit consumers' demand for online transactions. Preferences for face-to-face interactions, privacy concerns about transmitting certain information electronically, and the inability to touch, try on, or tangibly compare certain products online constrain participation in electronic commerce—even among people who already use the Internet. Telecommunications technology can contribute to easing those constraints, but is only one of several relevant factors. Network infrastructure is more cen-

48. See, e.g., Sharon Linstedt, *Santa Shops on the Web; \$3.5 Billion in Online Sales Set This Season*, THE BUFFALO NEWS, Dec. 20, 1998, at 14C.

49. See Mark E. Plotkin, *How Traditional Companies Can Navigate the Web*, LEGAL TIMES, Mar. 1, 1998, at S32.

trally relevant to the transaction costs of exchanges consumers do undertake electronically, as well as to the ability of online merchants to expand the range of transactions consumers are willing to engage in on the Internet.

Basic, copper telephone links generally allow data to be retrieved at a rate of about 56 kbps (at best). At that speed, still images download slowly and video displays can take prohibitive amounts of time. For example, to download a 3.5 minute video clip through a standard 56 kbps modem takes more than 20 minutes.⁵⁰ Even with a fast ISDN line, which transmits at about 128 kbps, that clip takes 10 minutes to retrieve.⁵¹ Such time requirements restrict the ability and incentive of potential customers to retrieve useful or necessary product information and reduce the number of transactions for which they are willing to spend the necessary time on the Internet. Even if they are willing to retrieve slow-loading visual images from one or two sites, their ability to browse new sites and compare price and product offerings among online merchants is limited. Frustrated with the effort, some users will either buy from an established online seller, or buy the item at issue on their next trip to the (real) mall, as the world will not soon dispense with the necessity of some conventional shopping no matter how fast e-commerce expands.

At faster speeds, consumers will obviously be able to explore more sites and, perhaps more importantly, to obtain higher quality information about products—such as video and voice descriptions, and interactive responses—without prohibitive delay. A customer connected to the Internet through a 4 Mbps cable modem can download the above-mentioned 3.5 minute video in mere 20 seconds, making shorter product videos almost instantaneously viewable.⁵² And even the more modest 384 kbps DSL service becoming available in some areas would speed the download time to under 4 minutes.⁵³ With such connections, which allow access to more enhanced, interactive information, Internet users will likely engage not just in more transactions, but in more *kinds* of transactions as well.

B. The Sellers' Perspective: Reducing Barriers to Entry

Slow infrastructure speeds are also an impediment to sellers, particularly to new entrants into established lines of electronic commerce. A

50. See U.S. GOV'T WORKING GROUP ON ELECTRONIC COMMERCE, FIRST ANNUAL REPORT 25 (Nov. 1998).

51. See *id.*

52. See *id.*

53. See *id.*

customer with limited time can browse a certain number of sites turned up by a search. If speeds now available to most households stay the same, the fact that more relevant sites come online will not proportionally increase the number the customer can visit, but instead shift and divide customers among sites.

To illustrate the dimensions of the problem, consider the explosive growth of Internet resources. Five years ago there were a couple million Internet “hosts”—computers that store sources of information on the Internet—in the United States.⁵⁴ By 1998, more than 35 million Internet hosts were active world-wide, up from 20 million only six months earlier, and from fewer than 3 million in 1993.⁵⁵ In 1993, there was roughly one Internet host in the United States for every 200 Americans.⁵⁶ By 1997, the ratio had changed ten-fold, to 1 host for every 20 people—about one Internet host for every four American adults who use the Internet.⁵⁷ Assuming that access speeds and time spent online by individuals has grown less quickly than the number of sources available on the Internet—which is certainly the case—the increase in hosts means a user will search a decreasing proportion of sites relevant for a particular transaction.

Competition among sites will bring consumers some benefits even if they cannot browse more sites per online session than they do now. The market does most of the work for them; prices will decline at the store one does go to because of competition from the store one has never shopped, but which other consumers patronize. New online marketplaces and sophisticated search tools that help consumers comparison shop make the market more effective at communicating prices. But a proliferation of new sites, without a significant increase in access speed, still means that a decreasing proportion of potential consumers may ever connect to any individual site, never mind choose to purchase from it.

The problem for the e-commerce market overall is that, at some point, the customer base divides sufficiently that retail entry becomes a poor prospect. In an environment where few purely online businesses are yet turning a profit, impediments to entry may have a non-trivial effect on the growth rate of electronic commerce. Several commentators suggest that it may already be too late for new entrants into the online market to succeed,

54. See PROGRESSIVE POL'Y INST., *supra* note 45, at 30.

55. See *id.*

56. See *id.*

57. See *id.*

and that at very least there will be a very high failure rate for new ventures.⁵⁸

Faster, cheaper Internet access will do three things to make entry into the electronic marketplace more attractive for online businesses. First, it will allow consumers to compare more sites in the time they allocate to online shopping, thus expanding the addressable market for competing businesses. Second, it will likely increase the number of transactions consumers choose to complete electronically by making them easier and more convenient when compared to alternative, non-electronic means. And third, broadband connections will help Internet businesses to expand the types of transactions consumers are willing to make online by supporting real-time interactive capabilities as well as voice, video, and other displays that increase the tangibility of products and services being examined electronically.

There are certainly ways to improve Internet applications, and e-commerce in particular, without changing telecommunications infrastructure. "Accelerator" programs are now available that will, while downloading information from a site, simultaneously load all sites linked to that site and speed return trips to sites already visited.⁵⁹ Shopping portals, sites that act as digital malls by organizing products or merchants into easily searched categories and allowing transactions to be paid for together at the portal's own "cash register," are convenient for buyers and provide sellers with a ready-made market.⁶⁰ These sites potentially cut search and transaction times. But in the end, infrastructure will still limit the speed with which sellers can be reached by potential customers.

Speed is particularly important to new entrants or to established sites that wish to launch new product offerings. Observers have already noted the phenomenon—called by one commentator the "killer click"—by which consumers' simple, initial choices have a long-term effect on competition in the online marketplace.⁶¹ This form of inertia, or perhaps path dependency, through which a bookmarked site becomes the default for the

58. See Bob Tedeschi, *Can Shopping Networks Survive a Crowded Market?*, N.Y. TIMES ON THE WEB, (Jan. 19, 1999) <<http://www.nytimes.com/library/tech/99/01/cyber/commerce/19commerce.html>> (citing various analyst comments and predictions).

59. See Gordon Bass, *Warp-Speed Web Surfing*, PC COMPUTING, Nov. 1998, at 128 (evaluating the performance of one accelerator program).

60. Examples are Yahoo's shopping area, 911gifts.com, and CyberShop. See, e.g., Tedeschi, *supra* note 58.

61. See George F. Colony, *My View: Killer Clicks* (visited Mar. 2, 1999) <<http://www.forrester.com/ER/Marketing/0,1053,61,00.html>>.

product or service at issue, is reinforced to the extent consumers find the potential benefits of competitors not worth the inconvenience of additional search time. To overcome the advantages that established online merchants have by virtue of their installed presence on consumer screens, new entrants into electronic commerce must first attract, and then hold, the attention of potential buyers.

The first can be accomplished through advertising, both on the web and in other media. Web advertising expenditures totaled about \$2 billion in 1998, more than double the level of 1997.⁶² But the second—holding the customer—can be more difficult, especially if bandwidth is a constraint. A buyer might log onto a new, online bookseller out of curiosity. But if exploring the site is slow, although no slower than the customer's bookmarked site, the customer is more likely to lose interest and revert to her familiar default site. If access is fast, however, the new entrant is more likely to be able to communicate the potential advantages of its site to the passing shopper and more likely to capture market share.

C. The Advantages of Broadband and the Challenge for Telecommunications

The commercial advantages of access speed make clear that the broadband deployment discussed in Part I will help the growth and competitiveness of electronic commerce. The fast growth of Internet usage will, of course, help e-commerce and other applications to grow regardless of data transmission rates. Between 1995 and 1997, the number of adults in the United States who used the Internet grew from about 14.3 million to over 41 million, or about one in five adults.⁶³ The latest Pew Center survey finds that the total number of American Internet users today is about 74 million.⁶⁴ But the growth from this increased usage will be all the greater with more widespread broadband deployment.

Given the importance for buyers and sellers of making advanced services rapidly and affordably available, regulatory and policy initiatives that can affect the path of such deployment are important for electronic commerce. The Federal Communications Commission ("FCC") is currently holding proceedings, pursuant to petitions filed under the Telecommuni-

62. See THE YANKEE GROUP, TVS, PCs, AND BEYOND: CONVERGENCE OR CONFUSION? (Dec. 1998).

63. See PROGRESSIVE POL'Y INST., *supra* note 45, at 31.

64. See Pew Research Center, *supra* note 47.

cations Act of 1996 ("the Act"),⁶⁵ that could potentially affect both the price and availability of residential broadband services in the coming years. The next section describes the proceedings and some potential concerns that those proceedings raise for electronic commerce.

III. REGULATION OF BROADBAND SERVICES AND IMPLICATIONS FOR E-COMMERCE

The above discussion of the broadband market shows that an enormous consumer market is yet to be served, several competitors are moving to serve it, and substantial benefits for e-commerce will result from such service. The questions that remain are the time frame, price and market conditions on which households will be able to purchase high-speed Internet access. The answers probably depend in large part on technological innovation, evolution of consumer demand, and the expanding range of services obtainable through broadband connections. But the answers will also be affected by regulation. The FCC is currently deciding how it will regulate broadband service offerings by telephone companies, with the stated goals of "encouraging the rapid deployment of new telecommunications technologies"⁶⁶ and "facilitat[ing] the ability of competing carriers to offer advanced services on equal footing with incumbent carriers and their affiliates."⁶⁷

This section will examine why the FCC's advanced services proceedings, despite having goals consonant with those of the electronic commerce industry, may lead to less rapid roll out than the unregulated market would provide. The explanations lie partly in the Act itself, and partly in the choice between preserving competition and allowing carriers to take full advantage of economies of scope that could potentially speed deployment through lower prices for consumers.

A. Background of the Advanced Services Proceedings

After the 1984 divestiture of AT&T, which broke up the integrated Bell System monopoly into a long distance company and seven separate,

65. Telecommunications Act of 1996, Pub. L. No. 104-104, 110 Stat. 56 (codified as amended in scattered sections of 47 U.S.C.).

66. *In re* Deployment of Wireline Services Offering Advanced Telecommunications Capability, FCC CC Docket No. 98-147 para. 1 (Aug. 6, 1998) [hereinafter Advanced Services Notice].

67. *Id.* para. 14.

local telephone companies,⁶⁸ local telephone service remained a franchise monopoly throughout the United States. The regional “baby bells” and GTE were, and remain, the largest local service providers, while over one thousand independent companies serve small, primarily rural, territories.⁶⁹ The regulatory barriers to entry into the local market were substantial and, for the most part, within the jurisdiction of state utilities commissions. Regulators have advanced a variety of justifications for exclusive local franchises: the economics of “natural monopoly,” preserving cross-subsidies that support universal service goals, and ensuring timely network upgrades and extensions. Competition was eventually allowed in the provision of “enhanced services,” like voice mail, but not generally in switched, local voice service.

The Telecommunications Act of 1996 radically changed that regulatory environment by preempting and prohibiting regulations that protect monopoly franchises for local telephone service.⁷⁰ The Act thus dismantled a legal and administrative structure that had evolved over decades and replaced it with the rule that local competition must be permitted. Moreover, the Act pushed this principle beyond the regulatory agencies to the incumbent local service monopolies themselves: it requires them to allow new competitors to interconnect⁷¹ to their networks and to lease elements of those networks necessary for the competitor to provide competing service, and to allow them to do so at cost.⁷² High-speed data service competitors have invested heavily in facilities, and several competitive DSL providers have entered multiple markets in which they compete against each other, the incumbent carriers, and cable modem providers. New DSL entrants in particular have taken advantage of the 1996 Act to lease customer loops⁷³ and rent space in the incumbent’s central office

68. See generally *U.S. v. AT&T*, 552 F.Supp 131 (D.D.C. 1982), *aff’d* 460 U.S. 1001 (1983).

69. See U. S. Telephone Ass’n, *USTA Fact Sheet* (visited Apr. 5, 1999) <<http://www.usta.org/ustafact.html>>.

70. See 47 U.S.C. § 253 (Supp. II 1994).

71. Interconnection means exchange of traffic between networks. If a new company entered the market, but its customers could not call, or be called by, customers of the established local company (e.g., Pacific Bell), then no one would subscribe to the new company. Interconnection allows the new entrant to provide the same network benefit as the incumbent.

72. See 47 U.S.C. § 251 (Supp. II 1994).

73. The customer loop is the most important network element, and the one that is most difficult for a new entrant to construct for itself.

(often referred to as "collocation") in order to offer their service to customers.⁷⁴

B. The FCC's Advanced Services Proceedings

The advanced services proceedings came about after several incumbent local carriers petitioned the FCC, pursuant to Section 706 of the 1996 Act, to allow them an exemption from the Act's local competition requirements for the provision of advanced services like DSL.⁷⁵ In other words, they want to provide high-speed services without having to allow competitors access to the unbundled elements used for those services or to sell those services at wholesale to new entrants that want to resell them. The FCC has declined to extend the complete forbearance sought by the incumbents.⁷⁶ But the Commission is considering allowing the incumbents to choose between two alternatives. The first would allow incumbents to provide advanced services free of the resale and unbundling requirements of the 1996 Act, but only if they did so through separate subsidiaries that dealt at arms length, and on the identical terms as outsiders, with the parent company (a safeguard often called "structural separation").⁷⁷ The second alternative would allow incumbents to provide advanced services directly, rather than through a subsidiary, but those advanced services and the facilities used to provide them would then be subject to the 1996 Act's resale and unbundling provisions—which means competitors would have access to these services and facilities at wholesale rates on an equal and non-discriminatory basis.⁷⁸

The rationale for the Commission's proposals is to preserve access to the local advanced services market for new competitors. Competitive local exchange carriers already may have difficulty getting the facilities they need from the incumbents, and therefore in gaining access to some customers lines. The problem is exacerbated by the fact that incumbent local service providers offer DSL service themselves at the same time that they control inputs—notably loops and collocation space—needed by their DSL competitors. The FCC's advanced services proceedings are designed

74. As discussed above, DSL technology uses special modems to transmit digital information over existing copper lines. "Collocation" allows competitors to place switching or other electronic equipment used to provide service over that loop in the Incumbent's central office.

75. See Advanced Services Notice, *supra* note 66, at para. 11.

76. See *id.*

77. See *id.*

78. See *id.*

to ensure that the incumbents do not discriminate against competitors in order to keep the broadband market for themselves.

C. The Regulatory Outlook: Increased Competition at the Cost of Cheaper Speed?

The FCC's market-opening goals may bring about a variety of competitive benefits over time. But in the specific context of advanced services, there may be a tradeoff between those competitive benefits and economies of scale and scope that could help consumers in the near term. Incumbent firms might, in their own words, be "uniquely well positioned among common carriers to bring advanced services to the mass market."⁷⁹ Indeed, under decades of operation as monopoly franchises, incumbent local carriers were able to construct extensive distribution networks serving the entire market, enabling them to realize scale economies and other advantages of incumbency that no other provider can yet match. Furthermore, the incumbents' ubiquitous local networks give them a platform from which other services (such as DSL) can be offered at only incremental cost, allowing incumbents to exploit economies of scope in the provision of multiple services.⁸⁰

Such economies of scope and scale create the potential for consumers to be served quickly and at lower cost by incumbents than by new entrants. It may be true that the incumbents' advantages stem from a history of regulatory protection and monopoly status, but that doesn't mean that these advantages are any less beneficial for consumers or desirable from the standpoint of electronic commerce. Yet two underlying constraints in the Commission's consideration of advanced services might prevent such efficiencies, assuming they exist, from being exploited by the incumbents, even if the incumbents face substantial competition in the advanced services market from cable or wireless companies.

The first constraint is rooted in the fact that under the Act, the FCC's primary mandate is to open markets to competition. The purpose of the advanced services proceedings is accordingly to "propose measures to promote the deployment of advanced services in a competitive manner by both local exchange carriers and new entrants."⁸¹ Competing carriers must

79. Comments of US West Communications, Inc., in FCC CC Docket No. 98-147, at 16 (Sept. 25, 1998), available at <<https://gullfoss.fcc.gov/cgi-bin/ws.exe/prod/ecfs/comsrch.htm>>.

80. See, e.g., *id.* at 16 ("The existence of extensive circuit-switched facilities will permit economies of scope in the roll-out of packet-switched technologies...").

81. Advanced Services Notice, *supra* note 66, para. 4.

be able to provide advanced services on "equal footing" with incumbents.⁸² Any exploitation of economies of scale or scope by the incumbents, regardless of the potential benefits for consumers, may therefore be blocked to the extent those economies give the incumbent an advantage over new competitors. This might not be the wrong decision over time. The benefits of competition are well established: it provides incentives to reduce prices, develop innovations, and improve quality. Without competition, efficiency gains could be kept by carriers as profits instead of being passed through to consumers. And one could perhaps argue that the long-term benefits of opening the market will exceed the near-term benefits of existing economies of scope and scale.

But there is a second constraint underlying the current proceedings that might cause efficiencies to be traded away even where competition exists. In its Memorandum Order, the Commission limits its competitive analysis to "wireline, broadband telecommunications services."⁸³ This limitation precludes the FCC from considering competition in the local market from non-telecommunications firms—namely cable providers.

Under the 1996 Act's definitions, cable is not a "telecommunications" service, and therefore does not factor into the FCC's analysis of whether the advanced services market is open to competition. The Act defines "telecommunications" as "transmission, between or among points specified by the user, of information of the user's choosing, without change in the form or content of the information as sent and received."⁸⁴ This definition fits neatly with voice or data transmission between parties, for which "common carriers" like phone companies provide mere conduit.⁸⁵ "Telecommunications" carriers neither choose nor alter content. Cable companies, by contrast, for the most part decide what will be transmitted over their networks. They choose and sell programming, not mere conduit, to customers. The 1984 Cable Act accordingly defines cable service as "one-way transmission to subscribers of (i) video programming, or (ii) other programming service,"⁸⁶ and establishes that cable providers are not

82. *Id.* para. 14.

83. *Id.* para. 3.

84. 47 U.S.C. § 153(43) (Supp. II 1994).

85. For an excellent discussion of the Act's definitional categories for regulation in the digital environment, see Jonathan Weinberg, *The Internet and "Telecommunications Services," Access Charges, Universal Service Mechanisms and Other Flotsam of the Regulatory System*, YALE J. ON REG. (forthcoming Spring 1999), available at <<http://www.msen.com/~weinberg/FLOTSAM.htm>>.

86. *Id.* § 522(6).

to be regulated as common carriers.⁸⁷ The statutory definition of cable service now includes “subscriber interaction, if any, which is required for the selection of such video programming or other programming service.”⁸⁸ This definition has been interpreted to include Internet access.⁸⁹

As a result of the statutory definition of “telecommunications” services, the substantial rivalry that incumbent telephone companies might face from cable systems in providing residential broadband services will not be considered in the Commission’s competitive assessment of the market for high-speed telecommunications. While such a constrained process might ensure there are more competitors in the broadband market, it does not ensure that there will be lower prices or even more vigorous competition. On one hand, the incumbents might be prevented from exploiting efficiencies that benefit consumers even though there is broadband competition from cable. On the other hand, cable providers will not face as strong competitive pressure from the incumbent phone companies if structural separation and non-discrimination obligations limit the incumbents’ ability to reduce prices. This outcome, whereby consumer benefits are lost unnecessarily, is one that regulators and Congress should strive to avoid.

The FCC’s discretion to consider the regulatory tradeoffs and to include cable or other non-“telecommunications” carriers in its analysis is limited by the Act. The Act’s provisions on advanced services do include regulatory forbearance as a method the FCC can consider to promote broadband deployment,⁹⁰ and moreover define advanced services “without regard to any transmission media or technology, as high-speed, switched, broadband telecommunications capability.”⁹¹

Even if the Commission could read that definition to include cable or new wireless technologies, and to decide that forbearance would best achieve fast deployment of broadband capability, it is unlikely that section 706 ultimately supplies independent authority for it to suspend regulation. Forbearance is defined and specifically governed by section 401.⁹² And while section 401 gives the Commission some statutory discretion to forbear from regulating incumbent telecommunications carriers, it does so

87. *See id.* § 541(c).

88. *Id.* § 522(6).

89. *See Speta, supra* note 10, at 52-55 (analyzing the statutory language and legislative history surrounding the definition of cable service).

90. *See* 47 U.S.C. § 706(a) (Supp. II 1994).

91. *Id.* § 706(c)(1).

92. *See generally id.* § 401.

only where regulation is unnecessary to ensure nondiscriminatory behavior towards competing carriers.⁹³ A finding that forbearance would serve the public interest and benefit consumers is necessary, but not sufficient under section 401, for the Commission to free incumbents from regulatory obligations.⁹⁴ Structural separation or other regulation may thus be driven solely by the Act's antidiscrimination mandate, even if such discrimination would be irrelevant—or even beneficial to consumers.⁹⁵

Although a full exploration of the possible statutory interpretations is beyond the scope of this essay, the foregoing analysis shows that using competition from cable or other non-"telecommunications" carriers as the basis for regulatory forbearance faces legal difficulties under the statute. Any remedy, if necessary, may in the end rest with Congress.

To make the case more concrete, consider the following example involving the joint provision of voice and high-speed data services over the same subscriber loop. Some DSL technologies can operate simultaneously with voice over the same line. Once a carrier is recovering its costs of operating that line from one service, the second service can be provided at incremental cost. As a general matter, however, most competitive data carriers entering the DSL market do not offer voice service, in part because the economics of voice delivery (which is subsidized for most residential customers) are very different from the economics of high speed data service. This puts new entrants at a potential disadvantage because a customer buying data service from the competitor must still, for the most part, buy voice service from the incumbent. Each charges a price that captures the line cost, so consumers pay for two lines. The incumbent, on the other hand, might be able to offer a lower total price when it provides both services itself: it can capture the line cost through one service and therefore be able to charge less for the second than if that service were purchased from a competitor on a stand-alone basis. Allowing firms to capture those economies, assuming they exist, could speed the deployment of advanced data services to residential customers.

To be sure, there may be debate over the existence and magnitude of scale and scope efficiencies. There may also be debate over whether those efficiencies need be lost as a result of regulation. But if economies of scope like the above do exist, then the FCC's two proposals for provision

93. *See id.* § 401(a)(1).

94. *See id.*

95. An additional hurdle is section 401(d), which prevents the Commission from forbearing until a carrier has "fully implemented" the Act's unbundling and interconnection requirements. *See id.* § 401(d).

of advanced services—either structural separation without unbundling obligations or no separation with unbundling obligations—may prevent economies of scope from being used by incumbents to underprice market entrants.⁹⁶ This may be good for competitors, but not necessarily for consumers, at least in the short run. There may also be ways to overcome the competitive disadvantages that come from the incumbents' network efficiencies without sacrificing the consumer benefits. Better collocation or joint marketing arrangements might be worked out that preserve scope and scale economies. But in the event they cannot be, regulators should be wary of sacrificing potential consumer benefits in the interests of preserving market entry. And policy makers should be especially hesitant to do so where, as in the broadband market, there is no natural monopoly and where a strong competitor exists that is not being factored into the regulatory analysis.

With cable, and perhaps eventually fixed wireless, in the market, incumbent telephone companies may be subject to competitive pressure to pass on efficiencies to consumers and invest in network innovation, regardless of whether there is competition on the residential telephone networks themselves.⁹⁷ Of course, duopoly (or triopoly) may not be ideal for purposes of either allocative efficiency or innovation incentives when compared to more competitive market structures.⁹⁸ But if the alternative to duopoly is regulation that encourages competitive entry by preventing exploitation of network efficiencies, then duopoly may be preferable. It may be best for consumers to let cable and telephone systems compete head-to-head, unregulated, in the broadband market.⁹⁹ Whether such a policy

96. SBC contends that more than half the costs of structural separation (amounting to some \$200 million) is attributable to the affiliate's "[p]urchasing separate loops for data use only without the proper authorization (certification) ... that would allow the affiliate to provide voice service as well as data service." *Ex Parte* Presentation of SBC Communications Inc., in CC Docket No. 98-147, at 1 (Nov. 20, 1998) (on file with author). In other words, the principal cost of separation, in SBC's view, stems from the fact that the affiliate cannot immediately offer voice and data services on an integrated basis.

97. It is also far from clear that competitors would be unable to survive in the event the incumbents exploited scale economies. This is especially so in the business services market, where product differentiation, an increased range of consumer options, and a different demand structure are likely to keep the market competitive.

98. See generally COUNCIL OF ECONOMIC ADVISORS, ECONOMIC REPORT OF THE PRESIDENT 171-218 (Feb. 1999) (discussing the affects of regulation on innovation).

99. Broadband regulation is at present asymmetric. Unlike incumbent telephone companies, cable companies do not have market-opening obligations under the 1996 Act, and the FCC has so far declined to consider separate "open access" requirements for ca-

would be wise depends on the substitutability of cable modem and DSL services and the relative cost structures for providing them. It also depends on an assessment of whether there are long-run benefits to competition within the DSL market itself that outweigh any scale and scope efficiencies from which consumers might benefit in the short run. The risk to consumers, and to electronic commerce, is that under the current statute it is difficult for policy makers to consider the foregoing questions in their regulatory analysis.

IV. CONCLUSION

Broadband deployment is increasing in the United States, and advanced telecommunications infrastructure is starting to become available to residential customers. Given the advantages of high-speed Internet access for electronic commerce, this is good news for online shoppers and merchants alike. Regulation of advanced services may, however, affect the speed of residential broadband deployment and the prices for such services in the short run. This essay does not try to resolve current regulatory questions about the proper competitive rules for the advanced services market. But it has tried to identify some important constraints underlying current proceedings, and to suggest that policy makers should be cautious not to let those constraints harm consumers, slow the expansion of affordable broadband services, or keep electronic commerce from reaching its potential rate of growth.

ble companies that enter the broadband services market. *See* Federal Communications Commission, Citing Pro-Competitive Benefits to Consumers, Commission Approves AT&T-TCI Merger, CS 99-2 (Feb. 17, 1999). There is intuitive appeal to the argument that if providers of one major broadband technology (cable modem) are not regulated, nor should the providers of the competing (DSL) technology be. For an argument supporting disparate regulation of cable and telephone systems in the broadband market, see Speta, *supra* note 23, at 58-80. Whether asymmetric regulation is warranted depends on a variety of factors, including the costs of regulation and the extent to which the disparately regulated carriers compete with each other.

STANDARDIZING GOVERNMENT STANDARD-SETTING POLICY FOR ELECTRONIC COMMERCE

By Mark A. Lemley[†]

ABSTRACT

The U.S. government's policy towards open standards in electronic commerce is inconsistent. On the one hand, the Magaziner Report endorses the idea of interoperable standards and open standard-setting processes for electronic commerce. It also suggests that governments should not be involved in setting technical standards. On the other hand, the Report also endorses government intervention in the standard-setting process in the case of encryption. Further, it recommends expanding intellectual property rights, without acknowledging the difficulties this can cause for open standards. Professor Lemley's article draws attention to this inconsistency, and suggests ways that the government could help promote open standards if it truly wished to do so.

The *Framework for Global Electronic Commerce* ("Magaziner Report" or "Report") contains strong language concerning the proper development of technological standards for electronic commerce.¹ Consistent with its general anti-government tenor, the Report takes a strong position against government standard-setting in its section on Technical Standards.² Somewhat more surprisingly, the Report also takes the position that technical standards should be open and promote interoperability,³ and

© 1999 Mark A. Lemley.

† Professor of Law, University of Texas School of Law; Of Counsel, Fish & Richardson, P.C. Thanks to Rose Hagan and the participants in the Berkeley Center for Law and Technology conference on the Legal and Policy Framework for Global Electronic Commerce for comments on an earlier draft.

1. WILLIAM J. CLINTON & ALBERT GORE, JR., A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE (1997), available at <<http://www.ecommerce.gov/framework.htm>> [hereinafter FRAMEWORK].

2. See *id.* § 9 ("The United States believes that the marketplace, not governments, should determine technical standards...."); see also *id.* at Background (referring to governmental control over standards development as a "[p]otential area ... of problematic regulation").

3. See *id.* § 9 ("Standards are critical ... as they can allow products and services from different vendors to work together"; "the marketplace ... should determine technical standards and other mechanisms for interoperability"; "Numerous private sector bodies

strongly suggests that the standards be set by industry groups rather than individual companies.⁴

Unfortunately, an examination of government policy towards electronic commerce reveals that the government's actual approach to standard-setting is internally inconsistent. Further, the broad general endorsement of open standards in the Report leaves a number of important issues unaddressed, including the role of intellectual property and antitrust law in standard setting. How these questions are dealt with in practice will have a significant impact on the way in which electronic commerce develops.

I. NETWORK EFFECTS AND THE NEED FOR STANDARDS

The Magaziner Report does not speak of standards for electronic commerce in a vacuum. Rather, it identifies a number of different areas where some sort of standard is necessary for electronic commerce to flourish. These areas include electronic payment systems (electronic funds transfers, e-cash, smart cards and the like), security infrastructure (encryption standards), contract infrastructure (standards for authentication, integrity, and non-repudiation), telecommunications, and data interchange.⁵

Virtually all of these issues arise in markets characterized by network effects,⁶ usually fairly strong ones. Thus, the interoperability of global telecommunications systems is obviously a precondition to international electronic commerce; if data from a buyer's system can never reach a seller's system, there can be no transaction. Similarly, electronic commerce between a particular buyer and a particular seller requires them to

have contributed to the process of developing voluntary standards that promote interoperability.”).

4. *See id.* (“[W]e urge industry driven multilateral fora to consider technical standards in this area.”). The Magaziner Report also endorses the standards model of the Internet Engineering Task Force (“IETF”), *see id.*; *see also, infra* Part III. On the other hand, the Report does note that “in some cases, multiple standards will compete for marketplace acceptance.” *See* FRAMEWORK, *supra* note 1, § 9.

5. *See* FRAMEWORK, *supra* note 1, § 9.

6. Network effects occur when the value to each consumer of a particular product is in part a function of how many other consumers buy that product. A good example is the telephone, which has no value to consumers unless it connects them to other people, preferably as many as possible. For a detailed discussion of network effects and law, see Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CALIF. L. REV. 479 (1998) [hereinafter Lemley & McGowan, *Network*].

agree on a method of payment, a method of assuring performance, a method of security, and a method of product delivery.⁷

It is possible, of course, to design such agreements for each transaction on an ad-hoc basis. Early electronic commerce has sometimes taken this form. In the absence of a widely-used email encryption program, for example, particular parties in an ongoing business relationship (say, lawyers and clients) will sometimes agree on an encryption mechanism for communications between them. But there are strong social benefits to having most parties use the same system, just as there are strong benefits to having most computer users work on the same operating system or word processing program. If everyone used the same encryption system or the same standards for data downloads, electronic commerce would be cheaper and easier than in a world without a dominant standard. The benefits are even more stark where there is little intrinsic difference between competing standards. There may be no inherent reason to prefer one brand of smart cards to another, for example, but there are certainly strong benefits to having a smart card that works in a large variety of vendors' machines.

In short, much of the infrastructure of electronic commerce would benefit from uniform standards. There are three basic ways to produce such standards. First, the government can mandate the choice of a particular standard, as happened historically in telecommunications and broadcasting. Second, industry players can come together in a standard-setting organization to select a single standard on which they will base their products. Finally, if the economic incentives for standardization are strong enough, no one needs to "choose" a standard at all: the market will "tip" to favor one particular product (a new, de facto standard) at the expense of competing products.⁸ Each approach has advantages and disadvantages.⁹

7. Parties can and do conduct electronic commerce without agreeing on all of these things expressly, but that simply means either that there is implicit acquiescence to a particular standard (the use of a credit card number, for example, or an "agreement" that the transaction will not be encrypted) or that the parties haven't considered the issue yet. For an argument that electronic cash has been unsuccessful largely because credit cards are being used online with great frequency, see John D. Muller, *Selected Developments in the Law of Cyberspace Payments*, 54 BUS. LAWYER 403, 406-07 (1998).

8. For a discussion of tipping, see Michael L. Katz & Carl Shapiro, *Network Externalities, Competition, and Compatibility*, 75 AM. ECON. REV. 424 (1985).

9. For a fuller explanation, see Mark A. Lemley, *Antitrust and the Internet Standardization Problem*, 28 CONN. L. REV. 1041 (1996) [hereinafter Lemley, *Standardization*]; Lemley & McGowan, *Networks*, *supra* note 6.

II. GOVERNMENT PARTICIPATION IN SETTING ELECTRONIC COMMERCE STANDARDS

The Magaziner Report takes a strong position against government-imposed standard setting for electronic commerce. It includes strong rhetoric on the importance of standards being set in the marketplace, not by a bureaucracy. This rhetoric seems to be directed abroad more than at home. Indeed, the First Annual Report of the U.S. government's Working Group on Electronic Commerce touts as a major accomplishment a resolution it pushed at the Global Standards Conference in 1997 in which government participants agreed to let the private sector lead in standard-setting.¹⁰ While there is talk of a limited role for government science agencies in facilitating development of open standards,¹¹ a reader of the Magaziner Report could be forgiven for thinking that the major project of the government in setting Internet standards will be to get out of the way.

Despite this strong rhetoric, the Administration has shown no hesitation in jumping into the standard-setting process when doing so would further its substantive goals. The most glaring example involves encryption policy. The government has for many years tried every means at its disposal, short of an outright ban, to prevent industry from coalescing around a strong encryption standard.¹² The Report claims that it is "encouraging the development of a voluntary, market-driven key management infrastructure," or at least that "in partnership with industry, [it] is taking steps to promote the development of market-driven standards."¹³ This is nonsense. For several years, the government has "encouraged" key escrow encryption by refusing to let industry export anything else¹⁴ and by refus-

10. U.S. GOV'T WORKING GROUP ON ELEC. COMMERCE, FIRST ANNUAL REPORT iv (Nov. 1998), available at <<http://www.doc.gov/ecommerce/E-comm.pdf>> [hereinafter FIRST ANNUAL REPORT].

11. See *id.* at 20 (discussing several facilitative roles for the National Institute of Standards and Technology).

12. For an early description of this fight, see A. Michael Froomkin, *The Metaphor is the Key: Cryptography, the Clipper Chip, and the Constitution*, 143 U. PA. L. REV. 709 (1995); see also A. Michael Froomkin, *It Came From Planet Clipper: The Battle Over Cryptographic Key "Escrow,"* 1996 U. CHI. LEGAL F. 15 (1996).

13. FRAMEWORK, *supra* note 1, § 6.

14. For cases discussing the government's regulation of encryption under the International Trafficking in Arms Regulations, see generally *Junger v. Daley*, 8 F. Supp. 2d 708 (N.D. Ohio 1998); *Bernstein v. U.S. Dept. of State*, 945 F. Supp. 1279 (N.D. Cal. 1996); *Karn v. U.S. Dept. of State*, 925 F. Supp. 1 (D.D.C. 1996).

ing to buy products that don't meet its idea of a proper standard.¹⁵ Even today, in a period of liberalization, the government still prohibits general commercial export of strong encryption products. One can agree or disagree with the government's arguments against strong encryption and in favor of a back door for law enforcement. I, for one, am largely agnostic on the subject. However, it is clear that in the area of encryption, the government has hardly left voluntary standard setting to take its own course.

Two other examples help illustrate this point. First, the Administration pushed Congress for years to enact some sort of criminal copyright statute regulating the development of so-called "circumvention devices" that facilitate copying.¹⁶ Indeed, its support for such a law is one of the centerpieces of the Magaziner Report.¹⁷ The Administration finally succeeded in 1998, when Congress passed the Digital Millennium Copyright Act ("DMCA").¹⁸ The primary purpose of the DMCA is to intervene in the innovation marketplace, by imposing what one might call "unilateral technological disarmament" on designers of encryption-breaking systems.¹⁹ Not only has the government once again intervened in the technological marketplace to promote its agenda, but it did so here in a way that is almost diametrically opposed to its efforts to push key escrow. One can perhaps discern a consistent government policy here—something along the lines of "the U.S. government should have the means to break encryption, but no one else should"—but it is hardly one consistent with the overarching principle that "the private sector should lead."²⁰

15. For a discussion of the government procurement policies related to key escrow encryption, see Howard S. Dakoff, Note, *The Clipper Chip Proposal: Deciphering the Unfounded Fears that Are Wrongfully Derailing Its Implementation*, 29 JOHN MARSHALL L. REV. 475, 482-84 (1996). Of course, supporting a standard in the marketplace is not the same as mandating choice of a particular technology. See Lemley & McGowan, *Networks*, *supra* note 6, at 544-45.

16. A circumvention device is one that effectively bypasses or disables a technological protection system designed to encrypt or restrict access to a piece of data, particularly a copyrighted work. See generally 17 U.S.C. § 1201(a)(1) (1998).

17. See FRAMEWORK, *supra* note 1, § 4.

18. Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2860 (1998).

19. For a superb analysis of this labyrinthine law, see Pamela Samuelson, *Intellectual Property and the Digital Economy: Why the Anti-Circumvention Regulations Need to be Revised*, 14 BERKELEY TECH. L.J. 519 (1999).

20. FRAMEWORK, *supra* note 1, § 1. Indeed, government policy in this area is sufficiently schizophrenic that Congress had to amend the original draft DMCA to provide an exception for law enforcement officials, who otherwise would be committing a crime by attempting to do what key escrow would allow them to do. See 17 U.S.C. § 1201(e) (1998).

Finally, it is worth considering the Report's approach to telecommunications regulation. True, the government has pushed quite strongly for telecommunications deregulation both here and abroad. But the Report also indicates that government telecommunications policy must be founded in part on "guaranteeing open access to networks on a non-discriminatory basis."²¹ I happen to think this is sound policy,²² and it does leave a good deal of room for private companies to maneuver. But a policy of open interconnection certainly contemplates government setting, if not the actual technical standards for telecommunications companies, at least the framework for those standards.

III. OPEN VS. CLOSED STANDARDS

The Magaziner Report's commitment to open, interoperable standards in telecommunications raises a more general issue for technical standard setting: the role of intellectual property in standards. Computer software is eligible for patent, copyright, trademark, and trade secret protection, as well as protection by contract.²³ While there is some debate over the scope of some of these rights as applied to industry standards, particularly copyright,²⁴ there seems no question today that a single company can at least

21. FRAMEWORK, *supra* note 1, § 7.

22. See Joseph Farrell, *Creating Local Competition*, 49 FED. COMM. L.J. 201, 211 (1996); Lemley & McGowan, *Networks*, *supra* note 6, at 551.

23. See generally PETER S. MENELL ET AL., LEGAL PROTECTION FOR COMPUTER TECHNOLOGY (forthcoming 1999).

24. Some courts, notably the First Circuit in *Lotus Dev. Corp. v. Borland Int'l*, 49 F.3d 807 (1st Cir. 1995), have held that standard protocols may be entirely ineligible for copyright protection. Even if an interface protocol is eligible for copyright protection, it may still lose protection against all but the most literal copying if a court concludes during its filtration analysis that virtually all of the elements of the interface are unprotectable. For example, in *Mitel v. Iqtel*, the Tenth Circuit concluded that the arbitrary selection of code numbers in the operation of telephone call controllers was not sufficiently original to qualify for copyright protection. See *Mitel v. Iqtel*, 124 F.3d 1366, 1373-74 (10th Cir. 1997). But see *American Dental Ass'n v. Delta Dental Plans*, 126 F.3d 977 (7th Cir. 1997) (concluding that the listing of code numbers assigned to each element was a copyrightable part of a taxonomy of information, because different numbers could have been chosen); *Atari Games Corp. v. Nintendo of Am.*, 975 F.2d 832, 840 (Fed. Cir. 1992) (holding that the arbitrary string of numbers in a lock-out device was not dictated by function and therefore could be copyrighted).

Program interface elements will also be uncopyrightable if they are dictated by external factors, such as the requirements of compatibility with a particular hardware or software platform. Thus, the Tenth Circuit concluded that the selection of values matched to codes in *Mitel's* call controllers, while sufficiently original to qualify for copyright protection, could not be protected because they were dictated by the needs of the indus-

own patent rights that dominate a standard²⁵ and probably copyright rights in a standard as well.²⁶ Recent developments make possible intellectual property ownership not merely of technical standards, but of entire business models in the electronic commerce environment.²⁷ The Magaziner Report generally endorses strengthening intellectual property rights in the electronic commerce environment.²⁸

It should be evident, however, that the goal of strengthening intellectual property protection is in some tension with the goal of promoting open standards for electronic commerce. While some intellectual property owners might choose to open their standards to competition—Sun Microsystems has done so with Java, for example²⁹—as a rule intellectual prop-

try. *Mitel*, 124 F.3d at 1366. See also *Computer Assoc. Int'l, Inc. v. Altai, Inc.*, 982 F.2d 693, 707-09 (2d Cir. 1992). On the other hand, other courts have found similar program elements to be copyrightable. See, e.g., *Atari*, 975 F.2d at 845; *Engineering Dynamics, Inc. v. Structural Software, Inc.*, 26 F.3d 1335, 1347 (5th Cir. 1994); *CMAX/Cleveland, Inc v. UCR, Inc.*, 804 F. Supp. 337, 355-56 (M.D. Ga. 1992).

25. For a general discussion of patents that confer market control on software standards, see Julie E. Cohen, *Reverse Engineering and the Rise of Electronic Vigilantism: Intellectual Property Implications of "Lock-Out" Technologies*, 68 S. CAL. L. REV. 1091 (1995).

26. For example, Microsoft's copyright in its operating system has served to give it effective control over the standards contained therein, despite questions as to whether Microsoft's applications program interfaces ("APIs") are themselves copyrightable. See Mark A. Lemley & David McGowan, *Could Java Change Everything? The Competitive Propriety of a Proprietary Standard*, 43 ANTITRUST BULL. 715 (1998) [hereinafter Lemley & McGowan, *Java*]. In part, this results from uncertainty about the copyrightability of the APIs themselves. In part, however, it also reflects the technical difficulty of designing a compatible operating system given the constraints of copyright law. See Lemley & McGowan, *Networks*, *supra* note 6, at 527-30.

27. See *State St. Bank & Trust Co. v. Signature Fin. Group*, 149 F.3d 1368 (Fed. Cir. 1998); see also U.S. Patent No. 5,794,210 (covering the concept of paying consumers to view ads); U.S. Patent No. 5,794,207 (covering buyer-priced auctions); U.S. Patent No. 5,790,793 (covering all Internet "push" technology); U.S. Patent No. 5,724,424 (covering a system of secure real time online payment); U.S. Patent No. 5,715,314 (covering electronic "shopping carts").

28. See FRAMEWORK, *supra* note 1, § 4 (endorsing the concept behind the DMCA, and encouraging patent protection for software and telecommunications inventions). See also *id.* § 3 (supporting the development of U.C.C. article 2B (now the draft Uniform Computer Information Transactions Act), which would effectively provide new and broader forms of intellectual property-like protection).

On the other hand, the Report does shy away from maximal protection in some areas: it is neutral on database protection, for example, and it encourages tougher reviews of software patent applications by the PTO. See *id.* § 4.

29. See Lemley & McGowan, *Java*, *supra* note 26, at 750-53.

erty ownership in a de facto standard is inimical to open standard setting.³⁰ The way to achieve a truly open, interoperable standard is to put the standard itself in the public domain. TCP/IP and HTML are good examples of public domain standards that nonetheless inspire both collaborative work to improve the standards (in the Internet Engineering Task Force ("IETF"), among other places) and the development of proprietary content for and extensions to the standards. One can imagine a world in which Microsoft owned the intellectual property rights in both TCP/IP and HTML, but it is hard to believe that the course of Internet development would have been the same.

Standard-setting organizations do offer a potential way to preserve open standards despite the presence of intellectual property. Many standard-setting organizations, including the IETF, have by-laws that restrict the ability of members to own or assert intellectual property rights in standards adopted by the group. These rules take a variety of forms. Some groups ban the ownership of intellectual property rights in standards altogether. Other groups may impose rules forbidding a member from asserting intellectual property rights in the standard altogether, or at least from asserting them against another member. Both of these approaches amount in effect to a royalty-free compulsory license. Still other groups allow members to retain intellectual property rights in a standard, but require that the intellectual property be licensed on "reasonable, nondiscriminatory terms" to those who wish to use the standard. Finally, some standard setting organizations merely require advance disclosure by members of any intellectual property rights that might cover a potential standard, so that the organization can use that information in deciding whether to adopt the proposed standard.

Such by-laws offer the possibility of preserving open standards even in a world of strong intellectual property rights. The by-laws are not without their problems, however. First, their enforceability is limited. Standard-setting organization by-laws obviously cannot bind companies who are not members of the organization; a company that is large enough (or that has a strong enough intellectual property portfolio) may simply choose to go it alone and develop a proprietary standard. Further, there is some question as to how internal rules will be enforced against members that violate those rules. Will agreeing to license a patent on nondiscriminatory terms

30. Indeed, even in the case of Java one might reasonably be concerned that if the standard prevails, Sun will assert its intellectual property rights in the standard to close it to others. *See id.* at 769-72 (raising this concern, and suggesting ways to deal with it).

actually bar a patent owner from filing suit for infringement? Will it limit her remedies under patent law? Or is such an agreement simply a contractual obligation enforceable only in a separate suit? Similarly, how can members of a standard-setting organization enforce a by-law that requires only disclosure of intellectual property rights, but not relinquishment or licensing of those rights? These are complex questions, and the law so far doesn't have entirely satisfactory answers.³¹ This uncertainty may make it hard to keep a standard open in the face of a determined effort by an intellectual property owner to close it.

Furthermore, the antitrust laws may restrict the rules standard-setting organizations can impose and enforce. A number of recent cases suggest that standard-setting organizations may not be free to compel members to license their intellectual property rights. One case, *Addamax v. Open Software Foundation*,³² went so far as to hold that the collective action of competitors in a standard-setting organization might itself violate the antitrust laws. While this case seems wrongly decided,³³ it may serve to deter standard-setting organizations from regulating member behavior at all. Further, standard-setting organizations that negotiate with patent owners on behalf of their members risk being characterized as a buyer's cartel trying to coerce a license at an artificially low price. The Antitrust Division of the Department of Justice has even taken action against the European Telecommunications Standards Institute for compelling members to relinquish claims of ownership in the standards it promulgates.³⁴ Finally,

31. Figuring out what the answers should be would take far more space than I have here. I hope to embark on this project in the near future.

32. 888 F. Supp. 274, 281, 284 (D. Mass. 1995).

33. See Lemley, *Standardization*, *supra* note 9, at 1080-90 (arguing that standard setting organizations should not be subject to antitrust liability in network markets unless they restrict rather than promote access to a standard).

34. In a series of negotiations regarding rules promulgated by the European Telecommunications Standards Institute ("ETSI"), the United States put substantial pressure on ETSI to back down from its original rule requiring disclosure and nondiscriminatory licensing of member intellectual property rights embodied in ETSI standards. For discussions of the evolution of ETSI intellectual property rules, see Cortien Prins & Martin Schiessl, *The New Telecommunications Standards Institute Policy: Conflicts Between Standardization and Intellectual Property Rights*, 8 EUR. INTELL. PROP. REV. 263 (1993); Mark Shurmer & Gary Lea, *Telecommunications Standardization and Intellectual Property Rights: A Fundamental Dilemma?*, in STANDARDS POLICY FOR INFORMATION INFRASTRUCTURE 391-96 (Brian Kahin & Janet Abbate eds., 1995).

Ironically enough, the Federal Trade Commission has taken the opposite position, bringing action against a member of a standard setting group that asserted patent rights in a group standard in violation of the organization's by-laws. See, e.g., *In re Dell*

the Federal Trade Commission and a private company recently filed actions against Intel alleging that Intel wrongfully retaliated against intellectual property owners that sued it by barring their access to Intel's own intellectual property.³⁵ While these cases don't directly concern standard setting organization rules, they may give such groups some hesitation about attempting to coerce compliance with their rules.

In short, we cannot be confident that standard-setting organizations can maintain open standards in the face of strong intellectual property rights governing those standards. As the number of companies that claim to own part or all of a standard increases, so does the likelihood that the standard produced by the marketplace will be owned by one competitor or another. The chance that a standard will truly be open correspondingly decreases.

IV. LESSONS FOR ELECTRONIC COMMERCE

It is certainly possible to overstate the importance of group standard setting to electronic commerce. Electronic commerce can occur without a universal open standard set by a standards body. Indeed, the phenomenal growth now occurring has taken place largely without the benefit of universal standards.³⁶

Computer Corp., No. 931-0097 (F.T.C. 1996). It seems odd to argue both that private group standard setting rules violate the antitrust laws, and that failing to comply with those rules violates the antitrust laws.

35. See *Intergraph Corp. v. Intel Corp.*, 3 F. Supp. 2d 1255 (N.D. Ala. 1998); *In re Intel Corp.*, FTC Docket 9288 (complaint filed June 8, 1998), available at <<http://www.ftc.gov/os/9806/intelfin.cmp.htm>>. The FTC case was recently settled by consent decree. Intel agreed not to cut off its supply of chips and technology to plaintiffs who sued it, provided they met certain conditions. See Federal Trade Commission, *In re Intel Corp.*, Agreement Containing Consent Order, FTC Docket 9288 (visited Apr. 14, 1999) <<http://www.ftc.gov/os/1999/9903/d09288intelagreement.htm>>.

36. This is not to say there are no such standards under development. See Mueller, *supra* note 7, at 412 ("In the midst of the vigorous competition among various companies and consortia to establish their proprietary payment methods, many leading financial services and technology companies are also participating in efforts to develop open technology standards."). Among these putative group standards for electronic commerce are BIPS, see Financial Services Technology Consortium, *The Bank Internet Payment System (BIPS): Leading the Way to Electronic Commerce* (visited Mar. 12, 1999) <<http://www.fstc.org/projects/bips/>>; the Account-Based Secure Payment Objects Standard, see CommerceNet, *Open Mailing Lists* (visited Mar. 12, 1999) <<http://www.commerce.net/resources/lists/open.html>>; the Internet Open Trading Protocol, see Internet Engineering Task Force, *Internet Open Trading Protocol* (last modified Feb. 16, 1999) <<http://www.ietf.org/html.charters/trade-charter.html>>; the Payment Fa-

Nonetheless, there seems no question that the growth of electronic commerce could be both faster and more efficient if a number of the infrastructure problems noted in the Magaziner Report were dealt with. A number of commentators in particular have noted that the widespread deployment of electronic payment systems—particularly smart cards and electronic cash—has been delayed by the lack of a single, interoperable standard for their use.³⁷ This is a classic network problem. Consumers won't invest in smart cards until they are widely accepted, and merchants won't accept them unless they expect consumers to use them. The problem is exacerbated by the fact that competitors' cards work on different standards, and so there is no guarantee that a given consumer's card will work with a given merchant's machine. The problem may be solved eventually if one competitor's card gains a dominant share of the market, but the uncertainty of the intervening period will delay widespread adoption of electronic commerce tools.³⁸ Further, the fact that one company owns the resulting standard may limit total market penetration of the standard in the medium run even after the company has won the de facto standards competition. Consumers and vendors locked into the losing standard, or who are prevented by exclusive dealing arrangements from dealing with the

cility Object Framework, see Object Management Group, *Electronic Payment RFP* (visited Mar. 12, 1999) <http://www.omg.org/schedule/Electronic_Payment_RFP.htm>; and the XML standard for extending HTML to enable machine-driven commerce, see Veo Systems, Abstract, *The XML Revolution in Internet Commerce* (visited Mar. 12, 1999) <http://www.veosys.com/xml/white_papers/whitepapers2.html>.

37. See, e.g., Elizabeth Judd, *BITS CEO Tells E-Money About the Group's Upcoming Initiatives*, E-MONEY, July 1998, at 26 (quoting Catherine Allen); Muller, *supra* note 7, at 410 ("Interoperability among different smart card systems is crucial to the development of smart cards, and industry leaders continue to work towards establishing worldwide open standards."); Richard Poynder, *Today's Technology: Understanding E-Money and E-Commerce*, E-MONEY, July 1998, at 18, 21; Cynthia Weaver, *Smartcards in the United States: What is Holding Up the Show?*, E-MONEY, Aug. 1998, at 3, 4.

38. See Weaver, *supra* note 37, at 4:

The worldwide problem of insufficient interoperability is another major hindrance to smartcard acceptance.... [The major vendors] are still waging a war over standards for their disparate electronic purse specifications.... Not until a universal standard for financial services applications is adopted by the international community will the United States forge ahead with smartcard open systems.

Id. See also Poynder, *supra* note 37, at 21 ("Until and unless a globally accepted purse architecture appears, or the main purse products are modified to be interoperable, then this situation [the prevalence of "proprietary, fragmented," standards that aren't interoperable] will do much to prevent the successful international proliferation of e-commerce....").

victor, will not be able to use the dominant standard.³⁹ In a network market, if significant participants are excluded from the standard, *everybody* loses.

The right choice between open and closed standards is a complex one. Factors that go into determining social welfare include the technical quality of the standards, the possibility of improving those standards over time, the variety and price of products that embody the standards, and the size of the market that will result, as well as the speed of adoption and durability of the winning standard. But in the context of electronic payment systems or transaction security, where the most important consideration is getting everyone speaking the same language, there are good reasons to think that open systems have a natural advantage. Certainly, the rhetoric of the Magaziner Report suggests the government is of that view.

If the government truly wants to promote open systems for electronic commerce, what should it do? Several possibilities come to mind. At a minimum, the government really should get out of the way of private standard setting organizations that promote open standards. The Administration's efforts to date have focused primarily on getting *foreign* governments not to intervene in the standard setting process.⁴⁰ But U.S. rules affect standard setting organizations as well. Antitrust challenges to organization rules that promote interoperability seem more likely to injure competition than to promote it. Also, the government should at least be aware that its efforts to block or alter a standard to achieve other policy goals—as it has done in the encryption context—will impose real costs on efforts to achieve a standard by consensus.⁴¹

But the government could take a more positive role in supporting the development of open standards. It could endorse interoperability in the marketplace in the same way it has endorsed key escrow: by refusing to buy or use products that rely on a closed proprietary standard. Alternatively, the government could simply require interoperability, particularly

39. Thus, in both the PC market and the VCR market, the result of a standards competition was that a minority of consumers and suppliers were excluded from the network even years after it was clear which standard had won the competition. *See, e.g.*, Paul David, *Clio and the Economics of QWERTY*, 75 AM. ECON. REV. 332, Issue 2 (May 1985) (noting the risk of such suboptimal lock-in).

40. *See* FIRST ANNUAL REPORT, *supra* note 10, at 20.

41. *Cf.* Jane Kaufman Winn, *Open Systems, Free Markets, and Regulation of Internet Commerce*, 72 TUL. L. REV. 1177 (1998) (arguing that government should not pass laws that favor particular technological choices before the technology is fully developed).

in markets that have historically been subject to regulation.⁴² The Magaziner Report itself endorses just such a requirement where telecommunications standards are concerned; perhaps the government's historic role in regulating currency and payment systems⁴³ could justify a similar requirement there. Finally, Congress or the courts could promote interoperability by erecting some limits on the scope of intellectual property protection, for example by precluding ownership of industry standards altogether or by permitting the copying of APIs where necessary to achieve interoperability. Copyright law has already made some strides in this direction;⁴⁴ it may be that patent law should contain such an exception as well.

42. Winn's caution against technology specific legislation, *see id.* at 1183, need not concern us overlong here. What I am suggesting is not a preference for a particular technical standard, but for a process of achieving that standard and for a particular set of rules regarding the use of whatever standard results.

43. For a discussion of the history of banking and payment system regulation in the United States, *see* Kerry Lynn Macintosh, *How to Encourage Global Electronic Commerce: The Case for Private Currencies on the Internet*, 11 HARV. J.L. & TECH. 733 (1998); Jane Kaufman Winn, *Clash of the Titans: Regulating the Competition Between Established and Emerging Electronic Payment Systems*, 14 BERKELEY TECH. L.J. 675 (1999).

44. Commentators have sharply divided on whether compatibility and/or standardization should justify reverse engineering or copying of parts of a plaintiff's computer program. Virtually all recent courts have endorsed reverse engineering in some circumstances. *See* Alcatel USA, Inc. v. DGI Technologies, Inc., 166 F.3d 772 (5th Cir. 1999); DSC Communications Corp. v. DGI Technologies, Inc., 81 F.3d 597, 601 (5th Cir. 1996); Bateman v. Mnemonics, Inc., 79 F.3d 1532, 1539 n.18 (11th Cir. 1996); Lotus Dev. Corp. v. Borland Int'l, Inc., 49 F.3d 807, 819-22 (1st Cir. 1995) (Boudin, J., concurring); Sega Enters., Ltd. v. Accolade, Inc., 977 F.2d 1510, 1527-28 (9th Cir. 1992); Atari Games Corp. v. Nintendo of Am., Inc., 975 F.2d 832, 843-44 (Fed. Cir. 1992); Vault Corp. v. Quaid Software Ltd., 847 F.2d 255, 270 (5th Cir. 1988); Mitel, Inc. v. Iqtel Inc., 896 F. Supp. 1050, 1056-57 (D. Colo. 1995), *aff'd on other grounds* 124 F.3d 1366 (10th Cir. 1997).

Most commentators have similarly endorsed such a reverse engineering right. *See, e.g.*, JONATHAN BAND & MASANOBU KATOH, *INTERFACES ON TRIAL* (1995); Julie E. Cohen, *Reverse Engineering and the Rise of Electronic Vigilantism: Intellectual Property Implications of "Lock-Out" Technologies*, 68 S. CAL. L. REV. 1091 (1995); Lawrence D. Graham & Richard O. Zerbe, Jr., *Economically Efficient Treatment of Computer Software: Reverse Engineering, Protection, and Disclosure*, 22 RUTGERS COMPUTER. & TECH. L.J. 61 (1996); Dennis S. Karjala, *Copyright Protection of Computer Documents, Reverse Engineering, and Professor Miller*, 19 U. DAYTON L. REV. 975, 1016-18 (1994); Maureen A. O'Rourke, *Drawing the Boundary Between Copyright and Contract: Copyright Preemption of Software License Terms*, 45 DUKE L.J. 479, 534 (1995); David A. Rice, *Sega and Beyond: A Beacon for Fair Use Analysis ... At Least as Far as It Goes*, 19 U. DAYTON L. REV. 1131, 1168 (1994).

Whether or not the government decides to stand behind the Magaziner Report's rhetoric in favor of open and private technical standards for electronic commerce, two things should be clear. First, the government is perfectly willing to intervene in the development of market standards when it has an interest in the outcome. Second, the choice between open and closed standards is an important one for the development of electronic commerce, and rhetoric alone won't produce the right outcome.

On the other hand, some early decisions rejected compatibility as a justification for copying. *See* *Apple Computer v. Franklin Computer*, 714 F.2d 1240 (3d Cir. 1983); *Digital Communications Assoc. v. Softclone Distributing Corp.*, 659 F. Supp. 449 (N.D. Ga. 1987). And one current case suggests limits on the reverse engineering right. *DSC Communications Corp. v. Pulse*, 1999 WL 126067 (Fed. Cir. 1999). *See also* Anthony Clapes, *Confessions of an Amicus Curiae: Technophobia, Law and Creativity in the Digital Arts*, 19 U. DAYTON L. REV. 903 (1994) (arguing that no right to reverse-engineer software should exist); Arthur Miller, *Copyright Protection for Computer Programs, Databases, and Computer-Generated Works: Is Anything New Since CONTU?*, 106 HARV. L. REV. 977 (1993) (same).

THE LIMITS IN OPEN CODE: REGULATORY STANDARDS AND THE FUTURE OF THE NET

By *Lawrence Lessig*[†]

ABSTRACT

This essay considers the effect of the open source software movement on government's ability to regulate the Net. Its claim is that an increase in open source software within the application space of the Internet decreases the government's power to regulate.

This is an essay about standards in the future of the Internet's governance. I begin with a distinction between two types of standards, and then continue with a reminder of a bit of history of the evolution of thought about regulation in cyberspace. I then draw upon this distinction and this history to suggest a question about the future of the Net's regulation. This question relates to the place of open source software in the future of the "application space" of the Internet. My argument is that open source software will make regulating cyberspace more difficult than it otherwise would be.

I. STANDARDS

Distinguish between two sorts of standards: coordinating and regulating. A coordinating standard is a rule that facilitates an activity that otherwise would not exist. A regulating standard restricts behavior within that activity, according to a policy set by the regulators. A coordinating standard can be imposed from the top down, or emerge from the bottom up; a regulating standard is ordinarily imposed only from the top down. Driving on the right side of the road is a coordinating standard. A speed limit is a regulating standard. Coordinating standards limit liberty (drive on the right) to make an activity possible (driving); regulating standards limit liberty within that activity (speeding) to advance a regulatory end (safety or

© 1999 Lawrence Lessig.

[†] Jack N. and Lillian R. Berkman Professor for Entrepreneurial Legal Studies, Harvard Law School. This essay is drawn from a lecture. Thanks to Susan Freiwald and Joel Reidenberg for pointing out the implications of this argument for regulation internationally. Thanks to Karen King for her research support.

fuel conservation). We understand why an individual would want to deviate from a regulating standard; it is (often) hard to make sense of a desire to deviate from a coordinating standard.

Standards on a computer network are similarly coordinating and regulating. TCP/IP is a coordinating standard—it is a convention that makes exchange of information over the Internet possible.¹ Space allocation on a network server is a regulating standard—it limits the storage space assigned to a particular user to allow many users to use the same storage resource.

Most of the important Internet standards to date have been coordinating standards—standards such as TCP/IP, FTP, and HTML. The Internet community has demonstrated well its ability to develop and deploy coordinating standards; this is the genius in organizations such as the Internet Engineering Task Force (“IETF”).² But in the future, most of the most significant debates about standards will be debates about regulating standards—about standards that allow the government to carry its policy choices into effect, whether or not those choices are the choices of bottom-up organizations like the IETF.

The Net’s success with standards in the future, then, depends upon the standards at stake. And its success with coordinating standards will not necessarily entail a similar success with regulatory standards.

II. REGULABILITY

That’s the distinction; now the history. It’s important that we remark how the debate about the regulation of cyberspace has changed. Three years ago the world was techno-libertarian. Frustrated sorts from our bureaucratic age looked to cyberspace as a place where regulation would not work, and hence as a place where people would be free. “Free” had two senses for these sorts—first, life in cyberspace was free from *any* regulation, and second, life there was free from regulation by *government*. Life

1. See generally Charles L. Hendrick, *Introduction to the Internet Protocols* (July 3, 1987) <<http://www.shiva.com/prod/techinfo/ip-intro.html>> (giving history as well as explanation of TCP/IP).

2. The IETF is the single most important Internet standards body, though it functions in a very different manner from ordinary standards bodies. Membership of the IETF is open, and standards get adopted only if implemented. See Internet Engineering Task Force, *Overview of the IETF*, (visited Apr. 1, 1999) <<http://www.ietf.org/overview.html>>. See also Scott Bradner, *The Internet Engineering Task Force*, in OPEN SOURCES: VOICES FROM THE OPEN SOURCE REVOLUTION 47 (Chris DiBona et al. eds., 1999).

in cyberspace, libertarians promised, was unregulated and *unregulable*. Behavior there was beyond the government's reach.

These were the ideas that defined first-generation thought about cyberspace and law. Law such as copyright was dead, lyricists such as John Perry Barlow sang.³ Law was fundamentally threatened, lawyers such as Post and Johnson warned.⁴ The Net would be a world where freedom reigned, and in some techno-Marxist way, governments would have no choice but to wither away.

These ideas did not go unchallenged. Rather, there were a few "crazies" around at the time who thought quite differently about regulation on the Net. I met two at a conference at Emory Law School three years ago, where they were busy challenging these then-commonplace ideas about the unregulated life of cyberspace.

One was then an assistant professor from Fordham: Joel Reidenberg. About the claim that life in cyberspace was free—unregulated at all—Reidenberg had a very different view. Life in cyberspace, Reidenberg argued, was regulated as any form of life was. This regulation, however, was built into the code.⁵ This form of regulation he called *lex informatica*,⁶ and this *lex*, he maintained, defines what behavior is possible in cyberspace and what values cyberspace will uphold.⁷ Whether these are values of anonymity or privacy or free speech or access, it is this law that makes those values possible.

But the *lex informatica*, he argued, was not a law that was fixed.⁸ The architectures of cyberspace could be changed. The values that cyberspace embraces could be different. There is no nature to the way that cyberspace

3. See, e.g., John Perry Barlow, *Keynote Address, Symposium on "Fundamental Rights on the Information Superhighway" at the New York University School of Law*, 1994 ANN. SURV. AM. L. 355 (1994); John Perry Barlow, *The Economy of Ideas*, WIRED, Mar. 1994, at 84.

4. See, e.g., David R. Johnson & David Post, *Law and Borders—The Rise of Law in Cyberspace*, 48 STAN. L. REV. 1367, 1375 (1996).

5. By "code" I mean generally the software and hardware that constitutes cyberspace as it is. That code might be divided between the basic net protocols of TCP/IP, and the applications that run on those protocols. As I explain more below, it is the application space that is the most important target of regulation.

6. See Joel R. Reidenberg, *Governing Networks and Rule-Making in Cyberspace*, 45 EMORY L.J. 911, 929 (1996). See also Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553 (1998) [hereinafter Reidenberg, *Lex Informatica*].

7. See Reidenberg, *Lex Informatica*, *supra* note 6, at 568-73.

8. See *id.* at 579-81.

is built—no nature, simply code. This code could be made to be very different from what it currently is. It could be made, that is, to embrace a very different set of values.

The other crazy was Pam Samuelson, then a professor at the University of Pittsburgh. Samuelson challenged the second idea—that cyberspace could not be regulated by government. For as Samuelson saw it, the law was already threatening an important regulation of life in cyberspace.⁹ Not directly, of course, but indirectly—through a series of changes threatened by the Administration's White Paper on Intellectual Property.¹⁰ These changes, designed to increase the law's protection for intellectual property, threatened to fundamentally queer the architectures of cyberspace. Laws would have their effect, if only indirectly, by inducing changes in the *lex* that Reidenberg spoke of.

Time works changes. The views of these two crazies have now become mainstream. Everyone now gets how the architecture of cyberspace is, in effect, a regulator. Everyone now understands that the freedom or control that one knows in cyberspace is a function of its code. Cookies¹¹ mean less privacy; choice about cookies means more privacy. A world without P3P¹² is a world with less control over privacy; a world with P3P is a world with more control over privacy. A world with PICS¹³ is a world where speech is less free; a world without PICS is, well, let's say, nice.¹⁴ The differences in these worlds are differences in the code of these worlds. Different code, different regulation, different worlds.

9. See, e.g., Pamela Samuelson, *Intellectual Property Issues Raised by the National Information Infrastructure*, 454 PLI/PAT 43 (1996).

10. Information Infrastructure Task Force, Working Group on Intellectual Property Rights, *Intellectual Property and the National Information Infrastructure: The Report of the Working Group on Intellectual Property Rights* (Sept. 1995) <<http://www.uspto.gov/web/offices/com/doc/ipnii/>>.

11. Cookies allow web sites to track users over multiple visits. See generally David Whalen, *The Unofficial Cookie FAQ, Version 2.51* (visited Apr. 1, 1999) <<http://www.cookiecentral.com/faq/index.shtml>>.

12. "The Platform for Privacy Preferences Project (P3P) enables Web sites to express their privacy practices and enables users to exercise preferences over those practices." World Wide Web Consortium, *Platform for Privacy Preferences (P3P) Syntax Specification* (working draft) (July 2, 1998) <<http://www.w3.org/TR/WD-P3P10-syntax-19980702>>.

13. The Platform for Internet Content Selection (PICS) is a protocol for facilitating the rating and filtering of content on the Internet. See World Wide Web Consortium, *Platform for Internet Content Selection (PICS)* (last modified Jan. 3, 1998) <<http://www.w3.org/pics>>.

14. See Lawrence Lessig, *Tyranny in the Infrastructure*, WIRED, Jul. 1997, at 96.

And so too do most now see how government might have a role in this regulation. Smart governments will regulate, but not by directly regulating the *behavior* of people in cyberspace. Smart governments will instead regulate by regulating the *code* that regulates the behavior of people in cyberspace. Cyberspace's code will become the target of regulation.¹⁵ The future will be littered with examples of government trying to intervene to assure that cyberspace is architected in a way to protect government's interests. Whether those interests will be interests against copyright management circumvention¹⁶ or interests in favor of encryption control,¹⁷ the government will increasingly see that the most efficient target of regulation is not people but binary code. Enslave the code while telling the world that you are leaving the space free¹⁸—this is the formula for taming the liberty that cyberspace now provides.

Two important conclusions follow from the arguments of these two crazies. First, if code is a kind of law, then we should focus, as we do with real-space law, on the freedoms and the constraints built into this code, and on how these freedoms and constraints are changing. And second, if governments regulate code, then we should think about the limits that should constrain government's power to regulate. For our constitutional tradition is one which limits governmental power by limiting government's direct legislative action; yet the future of the government's regulation of the Net is a future where government regulates by indirect legislative action. Constitutional values should constrain both indirect and direct regulation; so far it is not clear that they do.¹⁹

15. For a great example of regulation of the code, see the Oxley-Manton Amendment to the Security and Freedom Through Encryption (SAFE) Act, H.R. 695, 105th Cong. (1997), which would have regulated the type of permissible encryption to be just that which provided the required governmental access.

16. See Digital Millennium Copyright Act of 1998, Pub. L. No. 105-304, § 1201, 112 Stat. 2860, 2863-2872 (codified at 17 U.S.C. § 1201) (1998). For a critique of this anti-circumvention provision, see Pamela Samuelson, *Intellectual Property and the Digital Economy: Why the Anti-Circumvention Regulations Need to be Revised*, 14 BERKELEY TECH. L. J. 519 (1999).

17. See *supra* note 15.

18. See, e.g., WILLIAM J. CLINTON AND ALBERT GORE, JR., A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE (1997), available at <<http://www.iitf.nist.gov/eleccomm/ecom.htm>>.

19. My favorite example is *Rust v. Sullivan*, 500 U.S. 173 (1991), in which the Supreme Court upheld an indirect means of discouraging abortion, something the government cannot do directly. See Lawrence Lessig, *The New Chicago School*, 27 J. LEGAL STUD. 661, 670, 690-91 (1998).

III. LIMITS ON REGULABILITY

That's the history: Now something about the future. I want to focus on a new wrinkle in this debate about regulating cyberspace. We are just beginning to understand this new wrinkle, yet it may become the most important question about the future of cyberspace that we have yet seen.

You might think it follows from the commonplace views of our day—from those views once held by the crazies only, but now considered mainstream by most—that government is capable of effectively regulating the Net. If government can regulate the code, then government can require codewriters to build the standards that the government needs into the code. The future of regulatory standards under this view, then, would simply be a future where the government tells codewriters how to architect their code so as to incorporate governmental regulatory standards.

But in fact, the story is interestingly more complicated. In fact, this power of government depends upon a feature of the code—application space code²⁰—that has only recently become salient. This feature is its *ownership*. Whether government can regulate code depends in part upon who controls that code. If the code is closed—controlled by private for-profit organizations—then government's power is assured. But if the code is open—outside of the control of any particular private for-profit organization—then the government's power is threatened. The more application space code is open code, the less government can regulate that code.

The reason is straightforward. Open code is software in plain view. It is software that comes bundled with its source code as well as its object code. Object code is the code that the computer reads. If you display it on your machine, it will appear as gibberish. But source code is code that programmers can read.²¹ It is this code that allows a programmer to open an open source software project and see what makes it tick. By being able to see what makes it tick, open source software makes transparent any control that the code might carry. For example, if the code carries a government-mandated encryption routine, that routine will be apparent to open source coders. And because it is apparent, open source coders can then choose whether or not to adopt that portion of an open code project. For by its nature, and by the promises that it comes bundled with in the

20. I define this more *infra* in the text accompanying note 29.

21. Compilers can read it as well, but compilers simply turn this code from source code into object code.

form of licenses,²² any open code software project remains available for adopters to modify or improve, however the adopters think best.

Closed code functions differently. It does not come bundled with its source, which means that its code is hidden under a hood that won't open.²³ Thus adopters or users of closed code cannot as easily detect what makes closed code tick. They can't as easily see whether it carries within it a given encryption routine or systems for collecting private data or technologies for monitoring and reporting usage. Clever adopters can try to work it out through reverse engineering²⁴ or hacking. But no matter how clever the adopter, closed code will be harder to monitor, and harder to change than open code. An adopter of open source code who doesn't like a module can simply substitute another; an adopter of closed code has no equivalently simple choice.²⁵

This difference is critical to the question of regulability. For if the application space is built with closed code, then the ability of adopters to change that code is less than it would be if the application space were comprised of open code. If it is harder for adopters to change code, then it

22. See, e.g., Free Software Foundation, *GNU General Public License Version 2* (June 1991) <<http://www.gnu.ai.mit.edu/copyleft/gpl.html>>; Ira V. Heffan, Note, *Copyleft: Licensing Collaborative Works in the Digital Age*, 49 STAN. L. REV. 1487, 1508 (1997). ("The GNU GPL gives users permission to copy, modify, and distribute GNU software conditioned on the user's agreement to license all derivative versions under the same terms. Further, users must agree (1) not to establish proprietary rights in the software; (2) to provide the source code to anyone to whom they give the object code; (3) to include in the software notice of the applicability of the GNU GPL; and (4) to accept the software without warranties of any kind.") (footnotes omitted).

23. The idea is stolen (but can an idea be stolen?) from Austin Bunn, *Under the Hood*, FEED (visited Apr. 1, 1999) <<http://www.feedmag.com/oss/ossintro.html>>.

24. Though many licenses expressly forbid reverse engineering. See, e.g., Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CALIF. L. REV. 479, 528 (1998) ("Microsoft has argued that each of the 100 million-plus copies of object code it sells are limited distributions of trade secret information subject to a 'shrinkwrap license' agreement that prevents reverse engineering, and therefore that no one can obtain a copy of Microsoft's operating systems without 'agreeing' not to reverse engineer it."). I am with those who believe that copyright's rules about reverse engineering should be read to trump the contract promise to the contrary. See, e.g., Mark A. Lemley, *Beyond Preemption: The Law and Policy of Intellectual Property Licensing*, 87 CALIF. L. REV. 111 (1999); Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989 (1997).

25. Technically, this is misleading. Programmers of closed code do publish application program interfaces (APIs) that enable others to "plug in" to a closed application. In principle, if these were fully transparent, closed code would be closer to open code. The significant difference is that closed code APIs still cannot be modified.

is easier for governments to regulate through that code. Say the government has a standard it wishes to impose on some aspect of the application space. To the extent the regulatory standard gets imposed on closed code, it is more likely to be adopted by users than the same regulation imposed on open code. If the adopters don't like the regulatory standard (which, given the nature of many regulatory standards, is not unlikely), adopters can more easily swap out the regulated code if they use open code than if they use closed code.

An example offered by Peter Harter at this conference makes the point well. Netscape has turned its code for Netscape Communicator over to a version of the open source software movement. Its code is controlled by an organization called Mozilla, but its source is open. When Mozilla releases a new version, adopters around the world are permitted to download the source code, and adopt it or modify it as they wish.²⁶

The French government didn't get this idea. They wanted Netscape to modify the SSL standard²⁷ to enable decryption of SSL transactions, and so they asked Netscape to implement the request. But as Netscape reportedly told the French, there is really very little that Netscape can do to enable the cracking of SSL, and it is easy to see why.²⁸ Even if Netscape built a French version of SSL, enabling the French to spy whenever the French government wants, whether that version got *used* depends upon whether it is *adopted*. And even if Netscape put the French SSL version into the code of Netscape Communicator, there is no reason to expect that adopters of the code wouldn't simply substitute a different version of SSL for the French spy-enabled version. Whether the SSL code is adopted is a decision that rests with the users, not with Netscape.

26. See The Mozilla Organization, *Our Mission* (visited Apr. 1, 1999) <<http://www.mozilla.org/mission.html>> (announcing that Netscape Communicator and its source code would be available free of charge, and describing mozilla.org's role as a "clearing-house for the newly-available Netscape source ... to collect changes, help authors synchronize their work, and periodically make new source releases which incorporate the best work of the net as a whole").

27. SSL, the Secure Sockets Layer, is a security protocol developed by Netscape which "provides data encryption, server authentication, message integrity, and optional client authentication for a TCP/IP connection." Netscape Communications Corp., *Secure Sockets Layer* (visited Apr. 7, 1999) <<http://home.netscape.com/security/techbriefs/ssl.html>>.

28. Technically, there are two reasons why there is little that Netscape can do here. One is that SSL is an open standard, which Netscape doesn't control. But the second is the reason I am focusing on here: Even if it could control it, its "control" depends upon whether the code is open or closed.

Harter's example is an instance of my more general point: to the extent that code remains open, it is harder for government to regulate; to the extent it is closed, it is easier. Had the French demanded a change in a part of Netscape's code before Netscape had given its code to Mozilla, then it would have been much harder for adopters to identify and disable that code. But after the code is in the commons, governments' power is less. Thus my point: the regulability of the application space turns in part on whether the application space is open.

That's the claim, but it requires some qualifications.

First, my argument turns upon the nature of the "application space" code. This is not the distinction between operating systems and applications, but rather the distinction between the basic Internet protocols and the applications (or "ends") that depend upon or use these protocols. It is the design philosophy of the Net to keep the protocols simple and general, and to build sophistication and complexity into the ends.²⁹ It is possible to imagine the government trying to regulate the Internet's basic protocols. But because these are coordinating standards that effect very little substantive control on the content of the Net, they are unlikely to be the source of any powerful or significant regulation. Regulation, or regulatory standards, if they are to be effective, would have to be embedded in the application space code.

Second, my argument is not that a world with open code, or mostly open code, couldn't be regulated. In my view, there could be relatively small shifts in the architecture of the Net—in the functionality built into the application space—that would fundamentally enable state regulation, even if that application space were open code.³⁰ If the Internet became "certificate rich"—meaning that many people carried and used digital certificates³¹ while "on" the Net—local government's power to regulate the Net could fundamentally increase, whether or not the basic certificate architectures were open or closed code.

29. This design is more efficient, for building complexity into the protocols would not necessarily lead to simple ends. See, e.g., Jerome H. Saltzer et al., *End-to-End Arguments in System Design*, in INTEGRATED BROADBAND NETWORKS 30 (Amit Bhargava ed., 1991).

30. See Lawrence Lessig, *What Things Regulate Speech: CDA 2.0 vs. Filtering*, 38 JURIMETRICS J. 629 (1998).

31. Digital certificates "allow verification of the claim that a specific [encryption] key does in fact belong to a specific individual. Certificates help prevent someone from ... impersonat[ing] someone else." RSA Laboratories, *FAQ 4.0—Frequently Asked Questions About Today's Cryptography* (visited Apr. 7, 1999) <<http://www.rsa.com/rsalabs/faq/html/4-1-3-10.html>>.

Third, my argument is also not that a world with more closed code is always a world that is more regulable. Some closed code would not affect the Net's regulability. It matters little whether solitaire programs or certain utility programs are open or closed code, for there is little connection between them and any regulation the government might impose. (So long, that is, as they are as they say they are.) Thus the point about regulability is not a point about necessity; it is instead a point about possibility.

And finally, following from the third: my argument is not a criticism of closed code in general. I don't believe that the best possible world is one where all code is open, any more than I believe that the best possible real world is one where all property is public or part of the commons. There is a mix between open and closed spaces in real space and there should be a similar mix between open and closed spaces in cyberspace. The only enemy is the extremes—either a world that was perfectly propertized (either completely, or selectively if selected well), or a world that permitted no closed development. Whatever economic model might support projects like the GNU/Linux OS,³² there is no reason to believe the same model would work for every coding project.³³

* * * * *

To many in the open code movement, this whole argument about the values in open source software might seem quite odd. To them, the real issue with open source software is its power. Its real virtue is its amazing efficiency—its robustness and reliability. And no doubt, if these are its virtues, they are valuable indeed.

But my point is not to question any claim about efficiency. My point is simply that there are other issues at stake as well.³⁴ The architecture of cyberspace embeds a set of values, as it embeds or constitutes the possible. But beyond the values built into this architecture, there are values that are implicated by the ownership of code. Its ownership can enable a kind of

32. For a description of Linux, see Linux Online, *What is Linux* (last modified Mar. 16, 1999) <<http://www.linux.org/info/index.html>>.

33. See Brian Behlendorf, *Open Source as a Business Strategy*, in *OPEN SOURCES: VOICES FROM THE OPEN SOURCE REVOLUTION* 149 (Chris DiBona et al. eds., 1999).

34. This, I take it, is the strong and true point that Free Software Foundation founder Richard Stallman makes. See Richard Stallman, *Why Software Should Be Free* (Apr. 24, 1992) <<http://www.fsf.org/philosophy/shouldbefree.html>> (arguing that software ownership is harmful because fewer people use the program, none of the users can adapt or fix the program, and other developers cannot learn from the program, or base new work on it). See generally Free Software Foundation, *Philosophy of the GNU Project* (last modified Mar. 27, 1999) <<http://www.fsf.org/philosophy/philosophy.html>>.

check on government's power—a separation of powers that checks the extent that government can reach. Just as our Constitution embeds the values of the Bill of Rights while also embedding the protections of separation of powers,³⁵ so too should we think about the values that cyberspace embeds, as well as its structure.

However efficient open code may be, arguments about open source must also consider the questions that these values raise. For in my view, it makes as much sense to promote open source on efficiency grounds alone as it does to promote democracy on grounds of economic wealth alone. It may well be that democracies are more wealthy than other forms of government, just as it may well be that open source software is more robust than others. But it is a thin conception of value that would see wealth or efficiency as the only, or most important, value at stake.

35. See *Morrison v. Olson*, 487 U.S. 654, 710 (1988) (Scalia, J., dissenting) (“While the separation of powers may prevent us from righting every wrong, it does so in order to ensure that we do not lose liberty.”).

RESTORING AMERICANS' PRIVACY IN ELECTRONIC COMMERCE

By Joel R. Reidenberg[†]

ABSTRACT

In the United States today, substance abusers have greater privacy than web users and privacy has become the critical issue for the development of electronic commerce. Yet, the U.S. government's privacy policy relies on industry self-regulation rather than legal rights. This article argues that the theory of self-regulation has normative flaws and that public experience shows the failure of industry to implement fair information practices. Together the flawed theory and data scandals demonstrate the sophistry of U.S. policy. The article then examines the comprehensive legal rights approach to data protection that has been adopted by governments around the world, most notably in the European Union, but finds that difficulties implementing these laws for online services pose important challenges for the effective protection of citizens' privacy. The lessons show that safeguarding citizens' rights requires a combination of law and technology and that a legal incentive structure is necessary to stimulate the rapid development and implementation of privacy-protecting technologies. The article concludes with a recommendation for a framework privacy law in the United States modeled on the O.E.C.D. guidelines that includes a safe harbor provision for policies and technologies and that creates a U.S. Information Privacy Commission to assure the balance between citizens' privacy, industry needs, and global competitiveness.

Privacy is a critical issue for the growth of electronic commerce. During the last few years, an overwhelming majority of Americans report that they have lost control of their personal information and that current laws

© 1999 Joel R. Reidenberg.

[†] Professor of Law and Director of Graduate Program Academic Affairs, Fordham University School of Law. An earlier draft of this paper was presented at the University of California, Berkeley Symposium *The Legal and Policy Framework for Global Electronic Commerce: A Progress Report* held March 4-6, 1999. I am very grateful for the thoughtful comments of Symposium participants and of the editors of the *Berkeley Technology Law Journal*.

are not strong enough to protect their privacy.¹ In 1998, *Business Week* found that consumer worries about protecting privacy on the Internet ranked as "the top reason people are staying off the Web above cost, ease of use and annoying marketing messages."² The fair treatment of personal information and citizen confidence are each necessary conditions for electronic commerce over the next decade. Yet, sadly, at the political birth of the electronic commerce movement in 1997, the White House's report, *A Framework for Global Electronic Commerce*,³ more commonly referred to as the Magaziner Report, missed a key opportunity to assure the protection of citizens' privacy on the Internet.

For years, the United States has relied on narrow, ad hoc legal rights enacted in response to particular scandals involving abusive information practices.⁴ The approach has led to incoherence and significant gaps in the protection of citizens' privacy.⁵ For example, substance abusers have stronger privacy rights than web users in the United States.⁶ Yet, rather than revise American privacy protection, the Magaziner Report adopted a position enshrining the status quo.

This paper will first examine the philosophy and sophistry behind the U.S. policy of industry self-regulation. Next, the paper examines the com-

1. Privacy Exchange.org, *1998 Privacy Concerns & Consumer Choice Survey, Executive Summary*, at 1 (last modified Dec. 15, 1998) <<http://www.privacyexchange.org/jiss/surveys/1298execsum.html>> (reporting that 82% of those surveyed feel that consumers have lost all control over how companies collect and use their personal information); Am. Ass'n. of Retired Persons, *AARP Members' Concerns about Information Privacy*, Dec. 1998 (reporting that 78% of those polled found existing statutory protections inadequate to protect privacy).

2. *BW/Harris Poll: Online Insecurity*, *BUS. WK.*, Mar. 16, 1998, at 102 <<http://www.businessweek.com/1998/11/b3569107.htm>>.

3. WILLIAM J. CLINTON & ALBERT GORE, JR., *A FRAMEWORK FOR GLOBAL ELECTRONIC COMMERCE* (1997), available at <<http://www.iitf.nist.gov/eleccomm/ecommm.htm>> [hereinafter *FRAMEWORK*].

4. See PAUL M. SCHWARTZ & JOEL R. REIDENBERG, *DATA PRIVACY LAW: A STUDY OF UNITED STATES DATA PROTECTION* 10 (1996).

5. See generally FRED H. CATE, *PRIVACY IN THE INFORMATION AGE* (1997); SCHWARTZ & REIDENBERG, *DATA PRIVACY LAW*, *supra* note 4.

6. Federal law carefully protects the personal information of individuals who undergo treatment for alcohol or drug abuse in programs receiving federal funds or subject to federal regulation. See 42 U.S.C. §§ 290dd-1, 290dd-2 (1994); SCHWARTZ & REIDENBERG, *DATA PRIVACY LAW*, *supra* note 4, at 177-78. At the same time, only limited protection is available for Internet users. Statutory protection applies to telecommunications transaction information when collected by telecommunications service providers. See 47 U.S.C. § 222. However, if the data is collected by web sites, instead of service providers, then the statutory protection does not apply.

prehensive legal rights approach to data protection that has been adopted by governments elsewhere around the world, in a movement led by the European Union. While conceptually the cross-sectoral approach is better suited to the treatment of personal information in electronic commerce, the foreign experience illustrates a number of challenges for effective protection of citizens. The concluding section argues for a more desirable policy that combines legal and technological means in order to safeguard the privacy of citizens on the Internet.

I. THE PHILOSOPHY AND SOPHISTRY OF U.S. PRIVACY POLICY

Broad, international consensus exists on the basic standards of fair information practice and the protection of citizen privacy in a democratic society.⁷ As recently as June 1998, the Clinton Administration even said that the "O.E.C.D. Guidelines have served as the basis for virtually all privacy legislation and codes of conduct that have been developed over the years."⁸ Beginning with the U.S. Department of Health and Education's elaboration of the first computer privacy policy in 1973⁹ and the United States' approval of the Organization for Economic Co-Operation and Development's privacy guidelines in 1980, the United States has recognized benchmark norms for fair information practice. These norms include specification of the purpose for data collection, the consent of individuals to process personal information, the transparency of data processing, such as notice to individuals and access to their personal information, special

7. See Council of Europe, Convention for the protection of individuals with regard to automatic processing of personal data, Jan. 28, 1981, EUR. T.S. No. 108, *reprinted in* 20 I.L.M. 377 (1981), *available at* <<http://www.coe.fr/eng/legaltxt/108e.htm>> [hereinafter *European Convention*]; Directive 95/46/EC of the European Parliament and of the Council of 24 Oct. 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995 O.J. (L281) 31 (Nov. 23, 1995), *available at* <http://europa.eu.int/eur-lex/en/lif/dat/en_395L0046.html> [hereinafter *European Directive*]; O.E.C.D., RECOMMENDATIONS OF THE COUNCIL CONCERNING GUIDELINES GOVERNING THE PROTECTION OF PRIVACY AND TRANSBORDER FLOWS OF PERSONAL DATA, O.E.C.D. DOC. C58 (final) (Oct. 1, 1980), *reprinted in* 20 I.L.M. 422 (1981), *available at* <<http://www.oecd.org/dsti/sti/it/secur/prod/PRIV-EN.htm>> [hereinafter *OECD Guidelines*].

8. U.S. DEPT. OF COMMERCE, PRIVACY AND ELECTRONIC COMMERCE (June 1998) <<http://www.doc.gov/ecommerce/privacy.htm>>.

9. See U.S. DEP'T OF HEALTH, EDUC. & WELFARE, SECRETARY'S ADVISORY COMM. ON AUTOMATED PERSONAL DATA SYSTEMS, *Records, Computers and the Rights of Citizens* (1973), *reprinted in* U.S. PRIVACY PROTECTION STUDY COMMISSION, PERSONAL PRIVACY IN AN INFORMATION SOCIETY, 15 n.7 (1977).

treatment of particularly sensitive information, such as medical data, and the existence of enforcement remedies and mechanisms.

The United States, however, has rejected all attempts to legislate any full set of standards.¹⁰ Rather, Congress and state legislatures have enacted isolated and narrow statutes such as the Fair Credit Reporting Act¹¹ and the Video Privacy Protection Act,¹² after the discovery of particularly scandalous practices. This type of statutory protection only covers the particular activities committed by specific actors such as a consumer credit reporting agency or a video rental service provider. This reactive policy for fair information practices has historically been predicated on the philosophy that self-regulation will accomplish the most meaningful protection of privacy without intrusive government interference, and with the greatest flexibility for dynamically developing technologies. The theory holds that the marketplace will protect privacy because the fair treatment of personal information is valuable to consumers; in other words, industry will seek to protect personal information in order to gain consumer confidence and maximize profits.¹³

For more than twenty years, however, government agency task forces and reports regularly illustrated the lack of fair information practices in American society, but nevertheless resorted to the mantra that business should be given more time to self-regulate.¹⁴ With the Internet revolution,

10. See Robert M. Gellman, *Fragmented, Incomplete, and Discontinuous: The Failure of Federal Privacy Regulatory Proposals and Institutions*, 6 SOFTWARE L.J. 199 (1993).

11. 15 U.S.C. § 1681 (Supp. 3).

12. 18 U.S.C. § 2710-2711 (1994).

13. See, e.g., U.S. DEPT. OF COMMERCE, NAT'L TELECOMM. AND INFO. ADM., PRIVACY AND SELF-REGULATION IN THE INFORMATION AGE, Ch. 1.A (June 1997) <http://www.ntia.doc.gov/reports/privacy/privacy_rpt.htm>.

14. See, e.g., U.S. PRIVACY PROTECTION STUDY COMMISSION, PERSONAL PRIVACY IN AN INFORMATION SOCIETY (1977); FEDERAL TRADE COMMISSION, PRIVACY ONLINE: A REPORT TO CONGRESS (June 1998) <<http://www.ftc.gov/reports/privacy3/toc.htm>>; INFORMATION POLICY COMMITTEE, NATIONAL INFORMATION INFRASTRUCTURE TASK FORCE, OPTIONS FOR PROMOTING PRIVACY ON THE NATIONAL INFORMATION INFRASTRUCTURE (Apr. 1997) <<http://www.iitf.nist.gov/ipc/privacy.htm>>; FEDERAL TRADE COMMISSION, STAFF REPORT: PUBLIC WORKSHOP ON CONSUMER PRIVACY ON THE GLOBAL INFORMATION INFRASTRUCTURE (Dec. 1996) <<http://www.ftc.gov/reports/privacy/privacy1.htm>>; NAT'L TELECOMM. AND INFO. ADM., U.S. DEPT. OF COMMERCE, PRIVACY AND THE NII: SAFEGUARDING TELECOMMUNICATIONS-RELATED PERSONAL INFORMATION (Oct. 1995) <<http://www.ntia.doc.gov/ntiahome/privwhitepaper.html>>; U.S. ADVISORY COUNCIL, NATIONAL INFORMATION INFRASTRUCTURE, COMMON GROUND: FUNDAMENTAL PRINCIPLES FOR THE NATIONAL INFORMATION INFRASTRUCTURE (Mar. 1995) ; U.S. INFORMATION INFRASTRUCTURE TASK FORCE WORKING

the Clinton Administration had a chance to conceive a new vision of American privacy. Unfortunately for American citizens, the Magaziner Report sought to preserve the status quo:

The Administration considers data protection critically important. We believe that private efforts of industry working in cooperation with consumer groups are preferable to government regulation, but if effective privacy protection cannot be provided in this way, we will reevaluate this policy.¹⁵

In effect, the Magaziner Report catered to the industry of personal data rather than enshrining citizen participation in decisions about their personal data. Indeed, the marketplace of personal information is big business in the United States. By 1998, the gross annual revenue of companies selling personal information and profiles, largely without the knowledge or consent of the individuals concerned, was reportedly \$1.5 billion.¹⁶

Despite the claims of industry partisans, there are critical normative flaws in the theory of self-regulation for information practices. Self-regulation assumes that all privacy values can and should be resolved by a marketplace. Yet privacy interests are central to democratic governance¹⁷ and privacy has been hailed as a necessary condition for participatory governance.¹⁸ In contrast, totalitarian governments prefer the surveillance state.¹⁹ Indeed, a democratic government typically does not sell basic political rights. But even if one rejects this position, a marketplace can only function efficiently if there is transparency; citizens must be able to identify the collectors and users of their personal information. However, for personal information, the natural tendency of the marketplace is to obscure its treatment.

This is a classic case of market failure. Without disclosure by corporations, citizens cannot ascertain how their personal information is acquired and used. In the private sector, the economics are wrong for transpar-

GROUP ON PRIVACY, *PRIVACY AND THE NATIONAL INFORMATION INFRASTRUCTURE: PRINCIPLES FOR PROVIDING AND USING PERSONAL INFORMATION* (Oct. 1995) <http://www.iitf.nist.gov/ipc/ipc/ipc-pubs/niiprivprin_final.html>.

15. FRAMEWORK, *supra* note 3, at 14 (Issue 5).

16. See *In re Trans Union*, FTC Docket No. 9255, at 53 (July 31, 1998) <<http://www.ftc.gov/os/1998/9808/d9255pub.id.pdf>>.

17. See ALAN F. WESTIN, *PRIVACY AND FREEDOM* 23-26 (1967).

18. See Paul Schwartz, *Privacy and Participation: Personal Information and Public Sector Regulation in the United States*, 80 IOWA L. REV. 553 (1995); Spiros Simitis, *Reviewing Privacy in an Information Society*, 135 U. PA. L. REV. 707, 732 (1987).

19. See WESTIN, *supra* note 17, at 23.

ency.²⁰ Companies make significant profits from the secret collection and sale of personal information; the \$1.5 billion market in personal information is largely hidden from public view. Few individuals have ever heard of companies such as Acxiom or First Data. Yet, these companies have data warehouses with the most intimate details of the lives of millions of Americans. For example, Acxiom even sells information such as ethnic and religious affiliations, the type of car a person drives, and whether a person buys specialty clothing like particular types of underwear.²¹ Without transparency, an information trafficking industry has emerged in the United States with no accountability and minimal risk of harm to corporate financial interests from abuses of personal information. Not surprisingly, an analysis of industry codes of privacy practice reveals policies that fail to address the most basic principles of citizens' rights to personal information.²²

In effect, the American experience during the last two decades shows that the theory of self-regulation is pure sophistry. Time and again, the U.S. government has acknowledged that self-regulation remains hypothetical in corporate America. The Department of Commerce held a long awaited *Public Meeting on Internet Privacy* in June 1998, initially designed to give industry a chance to show its self-regulatory successes.²³ Unfortunately, industry had very little to show in terms of concrete implementation of privacy practices and the Secretary of Commerce conceded that the business community was failing to demonstrate effective self-regulation.²⁴ The Chairman of the Federal Trade Commission, in testimony to Congress during the summer of 1998, stated that "despite the Commission's considerable efforts to encourage and facilitate an effective

20. See Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1248 (1998) (observing that transaction costs are ignored in the market-based solutions); Paul Schwartz, *Privacy and the Economics of Personal Health Care Information*, 76 TEX. L. REV. 1 (1997).

21. See *Acxiom Catalog*, at 9 (ethnic data), 11 (specialty apparel data), 12-13 (car data) (1999) <<http://www.acxiom.com/infobase/catalog/catalog99.pdf>> (PDF file).

22. See Joel R. Reidenberg & Paul M. Schwartz, *Legal Perspectives on Privacy*, in INFORMATION PRIVACY: LOOKING FORWARD, LOOKING BACK (Mary Culnan & Robert Bies eds., forthcoming 1999) (noting particular failure of industry codes to encompass significant amounts of personal information and the failure to include remedies for victims of information abuse).

23. See U.S. DEPT. OF COMMERCE, *Agenda for Public Meeting on Internet Privacy* (June 23-24, 1998) <<http://www.ntia.doc.gov/ntiahome/privacy/confinfo/agenda.htm>>.

24. See Commerce Secretary William H. Daley, Remarks to Privacy Summit (June 23, 1998) (transcript available at <<http://www.doc.gov/opa/Speeches/980623.html>>).

self-regulatory system, we have not yet seen one emerge."²⁵ Several months later, the first government review of the position paper *A Framework for Global Electronic Commerce* wistfully admits that industry has only tentatively responded to privacy concerns even in the face of heavy government pressure.²⁶

It is worthy to note, however, that industry has improved its privacy talk over the last few years. Trade associations are now addressing the issues of data privacy (and lobbying Congress against regulation). The Secretary of Commerce has also tried to highlight self-regulatory initiatives such as TRUSTe and BBOnLine as evidence of progress.²⁷

But, ironically, these examples themselves demonstrate the structural defects in self-regulatory theory. TRUSTe, for example, is a program through which websites agree to disclose their privacy policies and license the right to use a special logo designating the site as one that protects privacy.²⁸ TRUSTe may audit licensees to verify compliance with the stated privacy policy. However, the program has had a few major problems. Although about 450 companies are licensed to use the logo to date, this number is trivial compared to the number of website operators in the United States. In fact, one of the companies, GeoCities, holds the distinction of being the first company prosecuted by the Federal Trade Commission for information trafficking,²⁹ and fifty percent of the TRUSTe sponsors do not bother to subscribe to the program and license the logo.³⁰ TRUSTe even features a link on its web page to a look-up service site that

25. *Electronic Commerce: Privacy in Cyberspace, Hearings on H.R. 2368 Before the Subcomm. on Telecommunications, Trade and Consumer Protection of the House Comm. on Commerce*, 105 Cong., 2nd Sess., July 21, 1998 (testimony of Robert Pitofsky, Chairman of the FTC), available at <http://www.ftc.gov/os/1998/9807/privac98.htm#N_3_>.

26. U.S. GOV'T WORKING GROUP ON ELEC. COMMERCE, FIRST ANNUAL REPORT 16 (Nov. 1998), available at <<http://www.doc.gov/ecommerce/E-comm.pdf>>.

27. See Commerce Secretary William H. Daley, Remarks at Press Conference on E-Commerce (Feb. 5, 1999) (transcript available at <<http://www.doc.gov/opa/Speeches/ecommerceremarks.html>>).

28. See TRUSTe, *TRUSTe Program Principles* (visited Mar.30, 1999) <http://www.truste.org/webpublishers/pub_principles.html>.

29. See *In re GeoCities Decision and Order*, F.T.C. Docket No. C-3850 (visited Mar.29, 1999) <<http://www.ftc.gov/os/1999/9902/9823015d&o.htm>>.

30. As of March 2, 1999, TRUSTe had 51 sponsors; only 26 were registered as licensees of the TRUSTe logo to show a commitment to privacy. Compare TRUSTe, *TRUSTe Sponsors* (visited Mar. 30, 1999) <http://www.truste.org/about/about_sponsors.html>, with TRUSTe, *Look Up A Company* (visited Mar.30, 1999) <http://www.truste.org/users/users_lookup.html>.

fails to disclose its privacy policy and is owned by a company that is not even listed as a TRUSTe licensee.³¹

A similar pattern exists at BBBOOnline, a project of the Better Business Bureau proposed more than a year ago in response to U.S. government pressure on industry to demonstrate that self-regulation might work.³² BBBOOnline hopes to provide an enforcement mechanism for privacy disputes online. However, for the moment, the BBBOOnline mechanism remains hypothetical. While the program officially launched on March 17, 1999,³³ BBBOOnline ignores the issue that consent might not be an appropriate basis for the processing of some personal information, such as health data, only requires that websites disclose particular practices, fails to require that remedies be afforded to victims of information abuse, and fails to require that individuals be granted complete access to their personal information.³⁴ In addition, BBBOOnline uses a nebulous and undefined term, "individually identifiable information," to circumscribe the scope of its participants' obligations. It also remains to be seen whether the online industry will participate on significant scale.

Another important privacy initiative likewise remains unavailable even after three years of development and government encouragement. Internet labeling and filtering technology based on the world wide web's protocol, Platform for Internet Content Selection ("PICS,") has been under development for a privacy application, the Platform for Privacy Preferences

31. TRUSTe requires that "web sites ... must disclose their personal information collection and privacy practices." TRUSTe, *The TRUSTe Program: How it Protects Your Privacy* (visited Mar. 30, 1999) <http://www.truste.org/users/users_how.html>. However, from the main TRUSTe member directory web page, TRUSTe, *Member Directory* (visited Mar. 30, 1999) <<http://www.truste.com>>, there is a link to <<http://www.worldpages.com/whitepages>>. This latter site allows a user to search for the address and phone number of anyone in the United States. The site does not display a TRUSTe logo, nor does it disclose any privacy policy. There is a link in fine print at the bottom of the web page *About Worldpages* to another web page: <<http://www.worldpages.com/docs/about.whml>> (visited Mar. 30, 1999). This last web page similarly says nothing about privacy, but does identify the owner of the page: Web YP, Inc. Web YP, Inc. is not listed as a licensee of TRUSTe, though a company identified as "World Pages, Inc." is listed.

32. See BBBOOnline, *Homepage* (visited Mar. 31, 1999) <<http://www.bbbonline.com>>.

33. See Robert O'Harrow, *Better Business Bureaus Offer Online Privacy Seal*, WASH. POST, Mar. 17, 1999, at E1.

34. See BBBOOnline, *Eligibility Criteria for BBBOOnline Privacy Seal* (visited Mar. 31, 1999) <<http://www.bbbonline.com/businesses/privacy/eligibility.html>>.

("P3P"), since 1996.³⁵ The World Wide Web Consortium ("W3C")³⁶, an influential standards setting body for the Internet, has led the development effort for P3P technology. Yet after three years, W3C has still not obtained sufficient industry agreement to conclude the development phase, let alone find companies willing to implement the technology. In addition, P3P faces a patent licensing problem that jeopardizes its ultimate adoption by industry.³⁷

The cornerstone of these self-regulatory efforts and U.S. policy seems to be the concept that notice and consent will solve the privacy issues. In describing the notice principle, the Magaziner Report articulates that "[d]ata-gatherers should inform consumers what information they are collecting, and how they intend to use such data."³⁸ The report describes the consent standard by asserting that "[d]ata gatherers should provide consumers with a meaningful way to limit use and re-use of personal information."³⁹ The Magaziner Report even argues that "principles of fair information practice [] rest on the fundamental precepts of awareness and choice."⁴⁰ This position is also emphasized clearly in the U.S. Department of Commerce's *Elements of Effective Self-Regulation*.⁴¹ Yet, these pronouncements seriously misconstrue basic fair information practices principles. These basic principles include key standards, such as purpose limitations, data minimization, and duration of storage that are not satisfied merely through notice and consent; notice and consent are not enough. The United States has even recognized this broader range of issues when it endorsed the O.E.C.D. Guidelines.⁴² In the rare instance when a government agency, the Federal Communications Commission,

35. See FEDERAL TRADE COMMISSION, TRANSCRIPT: PUBLIC WORKSHOP ON CONSUMER PRIVACY ON THE GLOBAL INFORMATION INFRASTRUCTURE, F.T.C. PROJECT P954807, at 79-90 (June 4, 1996) (statement of Paul Resnick, AT&T Research) (transcript available at <<http://www.ftc.gov/bcp/privacy/wkshp96/pw960604.pdf>>).

36. See W3C, *About the World Wide Web Consortium* (visited Apr. 20, 1999) <<http://www.w3.org/Consortium/>>.

37. See Chris Oakes, *Patent May Threaten E-Privacy*, WIRED, Nov. 11, 1998, available at <<http://www.wired.com/news/news/technology/story/16180.html>>; InterMind, *About InterMind Communication's Patents* (visited Apr. 20, 1999) <http://www.intermind.com/materials/patent_desc.html>.

38. FRAMEWORK, *supra* note 3, at 12 (Issue 5).

39. *Id.*

40. *Id.*

41. See U.S. DEPT. OF COMMERCE, N.T.I.A., ELEMENTS OF EFFECTIVE SELF-REGULATION FOR PROTECTION OF PRIVACY (Jan. 1998) <<http://www.ntia.doc.gov/reports/Elements/privacydraft/198dftprin.htm>>.

42. See *supra* note 8 and accompanying text; Gellman, *supra* note 10, at 200.

gave considered analysis to the effectiveness of consent as a legitimate basis for the sale of personal information to marketers, the FCC found opt-out to be a deficient basis for processing personal information under the Telecommunications Act of 1996 that mandated the protection of subscriber privacy.⁴³

Thus, to rely principally on notice and consent ignores the other basic fair information practice principles and underlines how self-regulation has not worked. Indeed, for the online world, technological defaults routinely favor privacy invasions over the implementation of fair information practices for citizens. Recent examples, such as the incorporation by Intel of an embedded identifier on each of its Pentium III chips⁴⁴ and the "smart browsing" features of Netscape Communicator and Internet Explorer software that upload from the user's computer a hidden file containing the Internet addresses of sites visited by the user,⁴⁵ illustrate techniques that facilitate the surreptitious surveillance of citizens. These examples demonstrate that the full range of fair information practice principles are marginalized by self-regulation defined in terms of notice and consent. Smart browsing, for instance, confronts the basic principle of purpose limitations and storage duration as addresses, processed to make website connections, are stored beyond the duration of the connection and now uploaded to a remote site for profiling purposes.

These basic flaws in the theory and practice of the U.S. self-regulatory approach pose an increasingly troubling problem for companies developing electronic commerce. Electronic commerce is global, yet American privacy policy is at odds with the growing movement around the world to establish clear, comprehensive legal rights. Ironically, American companies' global electronic commerce activities face an heretical choice: either provide better protection for U.S. citizens in order to have a single set of practices for global operations (because foreign laws require fair information practices) or maintain a double standard, treating foreign citizens to better privacy than U.S. citizens. The Magaziner Report largely ignores

43. See FCC Second Report and Order and Further Notice of Proposed Rulemaking, FCC Docket No. 96-149, ¶ 91 (Feb. 19, 1998) <http://www.fcc.gov/Bureaus/Common_Carrier/Orders/1998/fcc98027.txt>.

44. See Jeri Clausing, *After Intel Chip's Debut, Critics Step Up Attack*, N.Y. TIMES ON THE WEB (Feb. 19, 1999) <<http://www.nytimes.com/library/tech/99/02/cyber/articles/19intel.html>>.

45. See Netscape Corp., *What's Related FAQ* (visited Apr. 20, 1999) <<http://home.netscape.com/escapes/related/faq.html#o6>>.

this incongruity in boldly assuming that the rest of the world would simply accept the U.S. status quo with better educational efforts.⁴⁶

The international consequence of this self-regulatory pretense is an embarrassment for the U.S. government. Without demonstrable privacy protection in the United States, Europe threatens to block the flow of personal information to the United States.⁴⁷ The U.S. Department of Commerce has sought to negotiate with the European Commission a "safe harbor" code that would assure privacy for international data transfers to the United States and avoid any European data export prohibitions.⁴⁸ The proposal met with resounding criticism and virtual ridicule for its lack of content.⁴⁹ Because the Department of Commerce cannot propose any meaningful privacy standards, such as implementation mechanisms or enforcement devices providing remedies to victims, without undermining support for self-regulation, it is unequipped to respond to such criticism. Yet, without meaningful privacy standards, the United States isolates itself from the rest of the world. The time has come to reevaluate and reverse the policy that enshrines electronic surveillance and information trafficking against citizens.

II. THE CHALLENGE OF COMPREHENSIVE LEGAL STANDARDS

The recycling of unsuccessful and outdated privacy policies in the United States is in direct contrast to the data protection movement around

46. See FRAMEWORK, *supra* note 3, at 14 (Issue 5) ("The United States will continue policy discussions ... to increase understanding about the U.S. approach to privacy and to assure that the criteria [Europeans] use for evaluating adequacy are sufficiently flexible to accommodate our approach.").

47. See *European Directive*, *supra* note 7, at art. 25.

48. See U.S. Dept. of Commerce, *Draft International Safe Harbor Privacy Principles* (Nov. 4, 1998) <<http://www.ita.doc.gov/ecom/menu.htm>>.

49. See International Trade Administration, U.S. Dept. Of Commerce, *Public Comments filed on "Draft International Safe Harbor Privacy Principles"* <<http://www.ita.doc.gov/ecom/com.htm>>; Working Party of European Data Protection Supervisory Authorities, *Opinion 1/99 concerning the level of data protection in the United States and the ongoing discussion between the European Commission and the United States Government*, DG XV 5092/98/WP15 (Jan. 26, 1999) <<http://europa.eu.int/comm/dg15/en/media/dataprot/wpdocs/wp15en.htm>>; Working Party of European Data Protection Supervisory Authorities, *Transfers of personal data to third countries: Applying Articles 25 and 26 of the EU data protection directive*, DG XV D/5025/98/WP12 (July 24, 1998) <<http://europa.ue.int/comm/dg15/en/media/dataprot/wpdocs/wp12en.htm>>.

the world. Foreign countries, led by the fifteen states of the European Union (the "Member States"),⁵⁰ more typically follow an omnibus or comprehensive approach. Ironically, Europe learned its post-war lessons about information privacy from the movement in the United States during the 1960s and 1970s.⁵¹ But, unlike the United States, as European countries faced the computer processing of large quantities of personal information in the 1970s and 1980s, they adopted comprehensive data protection statutes to enshrine a rights-based, rather than market-based, approach to privacy. Indeed, in 1981, the Council of Europe opened for signature and ratification a data privacy treaty that has as its object and purpose "to secure in the territory of each Party for every individual, whatever his nationality or residence, respect for his rights and fundamental freedoms, and in particular his right to privacy, with regard to automatic processing of personal data."⁵²

Under the European model, framework legislation guarantees a broad set of rights to assure the fair treatment of personal information and the protection of citizens. In general, the modern European data protection laws define each citizen's basic legal right to "information self-determination."⁵³ This European premise of self-determination puts the citizen in control of the collection and use of personal information. The approach imposes responsibilities on data processors in connection with the acquisition, storage, use, and disclosure of personal information and, at the same time, accords citizens the right to consent to the processing of their personal information and the right to access stored personal data and have errors corrected. Rather than accord pre-eminence to business inter-

50. These states are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, and the United Kingdom.

51. See, e.g., COLIN BENNETT, *REGULATING PRIVACY: DATA PROTECTION AND PUBLIC POLICY IN EUROPE AND THE UNITED STATES* (1992); DAVID FLAHERTY, *PROTECTING PRIVACY IN SURVEILLANCE SOCIETIES* (1989); Fred H. Cate, *The EU Data Protection Directive, Information Privacy, and the Public Interest* 80 IOWA L. REV. 431 (1995).

52. *European Convention*, *supra* note 7, at art. 1.

53. This term "information self-determination" was coined by a 1983 German court decision prohibiting the intrusiveness of a national census. See Judgment of the First Senate [Bverfge, Karlsruhe], Dec. 15, 1983, *translated in* 5 HUM. RTS. L.J. 94 (1984).

ests, the European approach seeks to provide for a high level of protection for citizens.⁵⁴

Although the comprehensive rights approach has conceptual appeal for electronic commerce, it poses normative challenges for the structure of electronic commerce ventures and the effective protection of citizens. Because the rights-based approach relies on omnibus legislation, it covers the electronic processing of personal information regardless of context.⁵⁵ These statutes apply the same standards of fair treatment for personal information across sectoral boundaries of collection and use. In theory, this cross-sectoral application of principle correlates well to an information society where industry boundaries blur and data use defies clear categorization.

However, with the proliferation of European data protection laws during the course of the last two decades, the national laws evolved⁵⁶ and different standards in various Member States threatened the flow of personal information within Europe. For example, the scope of application of data protection laws and transparency requirements varied across national laws, posing conflicts for pan-European data processing.⁵⁷ In response, the Member States of the European Union sought to harmonize data protection principles and launched a five-year negotiating process that ultimately resulted in the enactment of the European Directive on data protection.⁵⁸

The European Directive confirmed the pre-existing comprehensive rights-based approach and contained both general and exacting rules aggregated from the laws of various European Union Member States.⁵⁹ Like the existing national laws, the European Directive's rules address the full set of internationally recognized principles. Each Member State must enact legislation implementing standards conforming to those defined by the

54. See, e.g., *European Directive*, *supra* note 7, at Recital 10 (explaining that the purpose of the Directive is to "seek to ensure a high level of protection in the Community").

55. See *id.*, at Recital 12, art. 3.

56. See Viktor Mayer-Schonberger, *Generational Development of Data Protection in Europe*, in *TECHNOLOGY AND PRIVACY: THE NEW LANDSCAPE* 220 (Philip E. Agre & Marc Rotenberg eds., 1998).

57. See *European Directive*, *supra* note 7, at Recital 7; JOEL R. REIDENBERG & PAUL M. SCHWARTZ, *DATA PROTECTION LAW AND ON-LINE SERVICES: REGULATORY RESPONSES* (Eur. Comm. 1998), available at <<http://europa.eu.int/comm/dg15/en/media/dataprot/studies/regul.pdf>>.

58. See *European Directive*, *supra* note 7.

59. See Spiros Simitis, *From the Market to the Polis: The EU Directive on the Protection of Personal Data*, 80 IOWA L. REV. 445 (1995).

European Directive,⁶⁰ and each Member State must maintain an independent, national supervisory authority for oversight and enforcement of these privacy protections.⁶¹ Significantly, the European Directive also mandates that Member State law require any person processing personal information to notify the supervisory authority and the supervisory authority must keep a public register of data processors.⁶²

While the harmonization of European data protection around comprehensive standards seems conceptually better suited to electronic commerce, in practice, the complexity of data processing arrangements in an information society makes the application of general principles to particular contexts challenging. Indeed, the registration mechanisms designed to assure transparency of processing activities can become onerous and problematic. Within Europe, critics have argued that compliance with these registration obligations is lacking.⁶³ Elsewhere, required notification to a government agency of data collection might be seen as an overly intrusive government action. In the United States, for example, the European commitment to the registration of data processing activities with a government agency would clash with Fourth Amendment values against government intrusion into the activities of citizens.

Furthermore, the application of the European Directive does not remove all divergences and ambiguities in the European national laws.⁶⁴ Small divergences and ambiguity will inevitably exist where the principles must be interpreted by different supervisory organizations in each of the Member States. These remaining divergences in standards can pose significant obstacles for the complex information processing arrangements typical in electronic commerce. For example, the European Directive requires that privacy rights attach to information about any "identifiable per-

60. This 'transposition' of the European Directive's standards into national law was to have occurred by October 1998. See *European Directive*, *supra* note 7, at art. 32. However, as is not uncommon in the European system, few Member States have complied with the deadline.

61. See *European Directive*, *supra* note 7, at art. 28.

62. See *id.* at art. 18-19.

63. See *Existing case-law on compliance with data protection laws and principles in the Member States of the European Union*, Annex to the Annual Report 1998 of the Working Party Established by Article 29 of Directive 95/46/EC (Douwe Korff ed., Eur. Comm: 1998).

64. See REIDENBERG & SCHWARTZ, *DATA PROTECTION LAW*, *supra* note 57; PETER SWIRE & ROBERT LITAN, *NONE OF YOUR BUSINESS: WORLD DATA FLOWS, ELECTRONIC COMMERCE, AND THE EUROPEAN PRIVACY DIRECTIVE* 188-96 (1998).

son.”⁶⁵ Yet, the scope of this definition is not the same across the Member States; what some Member States consider “identifiable” others do not.⁶⁶ Similarly, the disclosures that must be made to individuals prior to data collection vary within Europe.⁶⁷ These differences distort the ability and desirability of performing processing operations in various Member States since potentially conflicting requirements might apply to cross-border processing of personal information.

The effect of this challenge to comprehensive standards is, however, mitigated by consensus building options and extra-legal policy instruments that are available under the European model. The European Directive creates a working party of the Member States’ data protection commissioners.⁶⁸ The Working Party offers a formal channel for data protection officials to consult each other and to reach consensus on critical interpretive questions. But, policy guidelines from the Working Party will not be sufficient to assure privacy in electronic commerce. Guidelines will not be meaningful in a dynamic network environment without a technical infrastructure that also promotes data protection. This has been recognized internationally by data privacy commissioners: “it is mandatory to develop design principles for information and communications technology ... which will enable the individual user to control ... his personal data.”⁶⁹ Interestingly, the European model includes a provision for consensus on industry codes of conduct that might prove quite useful to facilitate the implementation of privacy compatible technologies.⁷⁰ The European Directive, building on Dutch law, provides for approval of codes of conduct as conforming to the privacy standards. This provision can be used to certify technical codes and configurations to assure privacy.⁷¹ The use of such technical measures may also be designed to avoid problems found in standards divergence, such as the differences in notice requirements.⁷²

65. *European Directive*, *supra* note 7, at art. 2(a).

66. *See* REIDENBERG & SCHWARTZ, *DATA PROTECTION LAW*, *supra* note 57, at 124-26.

67. *See id.* at 133-34.

68. *See European Directive*, *supra* note 7, at art. 29.

69. International Working Group on Data Protection and Telecommunications, *Data Protection and Privacy on the Internet: Report and Guidance* (Berlin, Nov. 18, 1996) <http://www.datenschutz-berlin.de/diskus/13_15.htm>.

70. *See European Directive*, *supra* note 7, at art. 27.

71. *See* REIDENBERG & SCHWARTZ, *DATA PROTECTION LAW*, *supra* note 57, at 147.

72. *See id.* at 153-54; Working Party of European Data Protection Supervisory Authorities, *Opinion 1/98: Platform for Privacy Preferences (P3P) and the Open Profiling Standard (OPS)*, DG XV D/5032/98/WP11 (June 16, 1998) <<http://europa.eu.int/>

For global information networks and electronic commerce, the comprehensive approach also inevitably invokes tension. Without the statutory authority to restrict transborder data flows, the balance of citizens' rights in Europe could easily be compromised by the circumvention of Europe for processing activities. Consequently, the European Directive includes two provisions to assure that personal information of European origin will be treated with European standards. The choice of law clause in the European Directive assures that the standards of the local state applies to activities within its jurisdiction and the transborder data flow provision prohibits the transfer of personal information to countries that do not have "adequate" privacy protection.⁷³ Some commentators have predicted that any European action will spark a trade war that Europe might lose before the new World Trade Organization.⁷⁴ While, in theory, such a situation is possible, it is equally remote.⁷⁵

Even with the difficulties of the European approach, countries elsewhere are looking at the European Directive as the basic model for information privacy, and significant legislative movements toward European-style data protection exist in Canada, South America, and Eastern Europe.⁷⁶ This movement can be attributed partly to the pressure from Europe arising from scrutiny of the adequacy of foreign privacy rights, but is also partly due to the conceptual appeal of a comprehensive set of data

comm/dg15/en/media/dataprot/wpdocs/wp11en.htm>; Joel R. Reidenberg, *International Data Flows and Methods to Strengthen International Co-operation* (paper presented at the 20th International Conference of Data Protection Authorities, Santiago de Compostela, Spain) (Sept. 17, 1998) <<http://home.sprynet.com/~reidenberg/idt.htm>>.

73. See *European Directive*, *supra* note 7, at art. 4 (choice of law) and art. 25 (export prohibition).

74. See SWIRE & LITAN, *supra* note 64, at 188-96.

75. See Joel R. Reidenberg, *The Movement toward Obligatory Standards for Fair Information Practices in the United States*, in *VISIONS FOR PRIVACY: POLICY CHOICES FOR THE DIGITAL AGE* (Colin Bennet & Rebecca Grant eds., 1999).

76. See, e.g., HUNGARIAN REPUBLIC, *THE FIRST THREE YEARS OF THE PARLIAMENTARY COMMISSIONER FOR DATA PROTECTION AND FREEDOM OF INFORMATION* 68-72 (1998) (discussing the influence of the European Directive for Hungarian data protection law); Council of Europe, *Chart of Signatories and Ratifications* (visited Apr. 20, 1999) <<http://www.coe.fr/tabconv/108t.htm>> (listing countries that have ratified the treaty on data privacy); Industry Canada, *Task Force on Electronic Commerce: The International Evolution of Data Protection* (Oct. 1, 1998) <<http://e-com.ic.gc.ca/english/fastfacts/43d10.htm>> (justifying the Canadian proposal for a comprehensive privacy law by reference to the European initiative); Office of the Privacy Commissioner for Personal Data, Hong Kong, *Personal Data (Privacy) Ordinance, Ch. 486* (visited Apr. 20, 1999) <http://www.pco.org.hk/ord/section_00.html> (displaying Hong Kong statute that follows the European comprehensive model).

protection standards. In effect, Europe has displaced the United States in setting the global privacy agenda with the enactment of the data privacy directive.

But, as illustrated by the European experience, the resolution of these difficulties cannot derive from law reform alone. In short, the comprehensive standards approach has two serious problems. First, general principles, while needed, leave significant margin for implementation and interpretation, especially in countries with very different legal cultures. For electronic commerce, any ostensibly small divergences in implementation or interpretation can generate significant distortions affecting the coverage for personal information and the incentives for protection by companies.⁷⁷ Second, the process to enact data protection law in Europe shows that adoption of legal rights is exceedingly slow. The existing European data protection directive took five years and transposition into national law was scheduled for three additional years.⁷⁸ In Internet time, these delays are generational.

III. SAFEGUARDING CITIZENS' RIGHTS WITH A COMBINATION OF LAW AND TECHNOLOGY

The lessons from the American experience with self-regulation show that government cannot abdicate responsibility for the protection of citizens' privacy to a marketplace skewed in favor of sale of personal information. At the same time, the lessons from the European experience involving detailed comprehensive statutes illustrate that effective privacy does not end with a legislative enactment. The guarantee of privacy for citizens requires a combination of law and technology that affords mechanisms to assure the fair treatment of personal information.

In a democratic state, privacy is and remains a basic right of citizens.⁷⁹ In contrast to many other aspects of privacy, informational privacy is unique in that citizens cannot determine how their personal information is being used without access to internal activities of those processing the data. To paraphrase Justice Stewart, "I do not know it when I cannot see

77. See REIDENBERG & SCHWARTZ, *DATA PROTECTION LAW*, *supra* note 57, at 142-46.

78. See *European Directive*, *supra* note 7, at art. 32.

79. See Jeb Rubinfeld, *The Right of Privacy*, 102 HARV. L. REV. 737 (1989); *OECD Guidelines*, *supra* note 7, at Preamble ("Member countries have a common interest in protecting privacy and individual liberties."); Schwartz *supra* note 18; Simitis, *supra* note 18; WESTIN, *supra* note 17.

it.”⁸⁰ As a consequence, the citizen confidence in the treatment of personal information that is so necessary for robust electronic commerce will not develop without a clear underlying set of rights.

To restore privacy for American citizens, the United States needs a framework that provides consistent fair information practices across different types of uses of personal information and different forms of processing arrangements. The United States government, however, need not try to reinvent fair information practice principles. The O.E.C.D. guidelines offer a full set of standards already recognized by the United States.⁸¹ The content of these guidelines provides a clear basis and level playing field for citizen privacy, and the guidelines themselves have been praised as sensitive to business concerns.⁸² These principles should be adopted in law as the American framework for information privacy.

Nevertheless, as both the American and European experiences show, technological capabilities and configurations hold the balance between effective fair treatment of personal information and defective privacy. Technical choices embed a set of policy rules for information flows in data processing systems. This “code”⁸³ or “lex informatica”⁸⁴ contained in the technical infrastructure has a direct rule-making effect on privacy. For ex-

80. *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1964) (describing attempts to categorize pornographic materials as “I know it when I see it.”).

81. See *O.E.C.D. Guidelines*, *supra* note 7; U.S. DEPT. OF COMMERCE, PRIVACY AND ELECTRONIC COMMERCE (June 1998) <<http://www.doc.gov/ecommerce/privacy.htm>> (recognizing the OECD Principles as the standard); U.S. Dept. of Comm., Nat’l Telecomm. and Info. Adm., *The Global Information Infrastructure: Agenda for Cooperation*, 60 Fed. Reg. 10359, 10367 (Feb. 24, 1995) (recognizing that the US accepts the OECD Principles).

82. After the O.E.C.D. adopted the guidelines, major U.S. companies subscribed to the principles. See GENERAL ACCOUNTING OFFICE, PRIVACY POLICY ACTIVITIES OF THE NATIONAL TELECOMMUNICATIONS AND INFORMATION AGENCY (Aug. 31, 1984) cited in Gellman, *supra* note 10, at 227 n.60; H.P. Gassman, *Vers un cadre juridique internationale pour l’informatique et autres nouvelles techniques de l’information*, ANNUAIRE FRANCAIS DE DROIT INTERNATIONAL 747, 750 (1985) (according to the author, who was a staff official at the O.E.C.D., 180 U.S. companies had subscribed to the O.E.C.D. guidelines).

83. See Lawrence Lessig, *Reading the Constitution in Cyberspace*, 45 EMORY L. J. 869, 898 (1996).

84. See Joel R. Reidenberg, *Governing Networks and Rule Making in Cyberspace*, 45 EMORY L. J. 911, 917-19, 929 (1996); Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules through Technology*, 76 TEX. L. REV. 553 (1998) [hereinafter *Lex Informatica*].

ample, the protocol P3P⁸⁵ is designed to empower web users by giving them information about website privacy policies and affording web users choices in the provision of personal information. However, P3P can only be effective if fairly written and appropriately implemented. The technical way in which the P3P protocol allows the expression of privacy policies and the choices given to web users are value-based decisions.⁸⁶ Furthermore, the manner in which P3P is incorporated in browsers, including the default settings and the fashion by which websites actually describe their practices, are critical for fair treatment of personal information. The development of "cookies" and their ability to track users across the Internet is another example of policy rules embedded in technical standards.⁸⁷ The initial default settings built into browsers encouraged the secret transfer of user's information, and only when faced with scandal did the software developers increase users' control over the disclosure of information.⁸⁸ These cases show that the technology can "go either way." The availability of privacy-protective technologies and privacy-enhancing default settings must exist. Yet, industry has demonstrated its lethargy in developing and implementing these technologies. Already, P3P has been in the development stage for three years and wide-spread use of the standard is, at best, a long time away.

Government must, therefore, act in a fashion that assures technological development in a direction favoring privacy protections rather than privacy intrusions. During the debate over self-regulation, U.S. industry took privacy more seriously only when government threats of regulation were perceived as credible. For example, the threats and cajoling from the Federal Trade Commission was a key impetus for the development of the BBBOOnLine, Online Privacy Alliance, and TRUSTe programs. But, despite deadline extensions for action by the Federal Trade Commission, none of these programs has yet to demonstrate accountability by their cor-

85. P3P is a protocol to enable disclosure and negotiation of the terms of consumer privacy between a web user and a web site collecting personal information. See W3C, *Platform for Privacy Preferences P3P Project* (visited Mar. 31, 1999) <<http://www.w3.org/P3P>>.

86. See Joel R. Reidenberg, *The Use of Technology to Assure Internet Privacy : Adapting Labels and Filters for Data Protection*, LEX ELECTRONICA (Fall 1997) <<http://www.lex-electronica.org/reidenbe.html>>.

87. See Mark Slayton, *An Introduction to Cookies*, HOT WIRED, Nov. 7, 1996 <<http://www.hotwired.com/webmonkey/webmonkey/geektalk/96/45/index3a.html>>.

88. See James Glave, *Next Netscape Will Chew Cookies on Command*, WIRED NEWS, Feb. 22, 1997, available at <<http://www.wired.com/news/news/technology/story/2196.html>>.

porate members for violations of privacy to individuals.⁸⁹ Indeed, to the contrary, industry created policies tend toward privacy myopia in the development of new products. Intel, for example, seemed genuinely surprised by the outrage expressed against its planned use of a unique identifier on its Pentium III chips.⁹⁰

With the enactment of a basic set of rights, the incentive structure for industry would shift to the development of effective protection for citizen privacy rather than the elaboration of vague policies to forestall corporate accountability. The existence of basic legal rights will force industry to deploy fair information practices that are well-balanced rather than skewed against citizens. To stimulate the quick development of privacy protecting system designs, these legal rights should allocate liability to companies that fail to develop and deploy privacy-enhancing technology.⁹¹ In doing this, legal standards will create new markets and opportunities for the development of privacy protecting products.

In any case, the promotion of privacy-friendly technologies and the implementation of fair information practices in particular contexts and especially in the electronic commerce context require constant vigilance. While counterintuitive for many in the United States, a U.S. Information Privacy Commission is urgently needed. Privacy policy requires a forum with a clear mandate for independent judgment to build consensus on solutions in particular contexts and to arbitrate disputes among stakeholders. In addition, U.S. business interests need an advocate in the face of international data flows. For years, the United States has remained on the sidelines of the annual meeting of data protection commissioners from around the world because the United States has no privacy commission.

At present, no existing agency or department in the United States is well suited to the tripartite role of consensus builder, privacy arbitrator, and international advocate. The Department of Commerce, where international privacy policy is presently formed, may be politically expedient, but is inappropriate for the range of privacy issues in the Information Society. The Commerce Department does not, for example, have particular expertise or competence in health privacy issues or global flows of employee data and is notoriously captured by business interests at the expense of

89. None of the programs offers any damage remedy to individuals when the company adherents fail to fulfill their privacy commitments.

90. See Polly Sprenger, *Intel on Privacy: 'Whoops!'*, WIRED NEWS, Jan. 25, 1999 <<http://www.wired.com/news/news/politics/story/17513.html>>.

91. See *Lex Informatica*, *supra* note 86, at 584 (discussing the effect of liability and the structure of the Internet.).

citizens' concerns.⁹² The State Department might be more appropriate for the foreign policy role, but has no expertise on the myriad of domestic privacy issues. Similarly, existing independent agencies such as the Federal Communications Commission would be poor choices for the centralization of privacy policy. The competence of these existing agencies is sectoral and each lacks expertise in cross-sectoral issues. The recent creation of a new position in the White House Office of Management and Budget is a good, but insufficient step.⁹³ Unfortunately, the new position is placed within the layers of the OMB bureaucracy and does not fulfill all the needed roles. Instead, the post has a coordinating role and does not have policy decision-making authority nor does the position have authority for the international negotiations with Europe.

If the United States hopes to protect effectively citizen privacy in electronic commerce, an independent privacy commission offers a number of attractive benefits both for citizens and businesses. The application of general privacy principles in the dynamic and complex online environment will inevitably require interpretation of the standards. Since a citizen's perspective may undervalue the interests of industry and society at large to information flows, while a corporate perspective will undervalue citizen's privacy, an independent privacy commission can offer critical guidance. In particular, such a commission can be accorded the authority to grant safe harbor protections for company practices.⁹⁴ Like a no-action letter from the Securities and Exchange Commission, a company seeking guidance and assurance that its policies are appropriate should be able to request approval from the privacy commission. Such an approval would mean that the practice conforms to the legal obligations for the fair treat-

92. For example, instead of publishing notice in the Federal Register for public comment on the draft international privacy principles, Undersecretary Aaron sent a letter, dated November 4, 1998, addressed "Dear Industry Representative" and posted it on a hidden web page several days later. See Letter from David Aaron, Undersecretary of Commerce to Industry Representatives (Nov. 4, 1998), available at <<http://www.ita.doc.gov/ecom/aaron114.html>>.

93. Declan McCullagh & James Glave, *Clinton Tabs Privacy Point Man*, WIRED NEWS, Mar. 3, 1999, available at <<http://www.wired.com/news/news/politics/story/18249.html>>.

94. See Joel R. Reidenberg, *Privacy in an Information Economy: A Fortress or Frontier for Individual Rights?*, 44 FED. COMM. L.J. 195, 242 (1992) (proposing a legislative model with a safe harbor mechanism for industry).

ment of personal information. This safe harbor approach was recently endorsed by the Federal Trade Commission.⁹⁵

In the context of electronic commerce, the safe harbor concept is especially powerful for guidance on technical infrastructure decisions. Technical protocols, default settings, and implementations can be treated the same way as company practices and policies for purposes of a safe harbor.⁹⁶ The existence of such a voluntary approval mechanism would give companies an important tool to avoid myopic, internal evaluations of the privacy ramifications, protect against data scandals, insulate the company from liability for privacy invasions, and satisfy foreign privacy regulators such as those in the European Union.

At the same time, the safe harbor process would afford citizens an opportunity for public comment on the conformity of practices to framework legal obligations and would not immunize practices outside the safe harbor nor immunize those safe harbor practices that change. Over time, safe harbor decisions would develop a body of public guidance that would increase transparency for all citizens. For citizens, the independent commission and a safe harbor procedure would also assure that the interpretation of fair information practices for electronic commerce continues as an ongoing process.

IV. CONCLUSION

The time has come for the U.S. government to become serious about privacy and restore protection to citizens. The Magaziner Report clearly erred in charting a conventional approach for a most unconventional, new environment. Citizens participating in global electronic commerce need to be assured that their personal information will be treated fairly. Companies engaged in electronic commerce cannot be crippled in their use of personal information. Fundamental values are at stake and one-sided policies and solutions will undermine democratic society.

95. See *Electronic Commerce: Privacy in Cyberspace, Hearings on H.R. 2368 Before the Subcomm. on Telecommunications, Trade and Consumer Protection of the House Comm. on Commerce, 105th Cong., 2nd Sess., July 21, 1998* (testimony of Robert Pitofsky, Chairman of the FTC), available at <http://www.ftc.gov/os/1998/9807/privac98.htm#N_3_>.

96. See, e.g., REIDENBERG & SCHWARTZ, *DATA PROTECTION LAW*, *supra* note 57, at 153-54.

ARTICLE

**DATABASE PROTECTION AT THE CROSSROADS:
RECENT DEVELOPMENTS AND THEIR IMPACT ON
SCIENCE AND TECHNOLOGY**

By J. H. Reichman and Paul F. Uhlir[†]

ABSTRACT

This article explores the potentially adverse impact that the emerging legal infrastructure could have on scientific, technical, and educational users of factual data and information—as well as on other sectors of the information economy—unless suitable adjustments are made. It begins by explaining how efforts to accommodate the networked environment to the publishers' fears of market failure will impose a daunting array of legal and contractual restraints on the ability of scientists and engineers to access factual data and information in the near future. It then goes on to examine the most recent efforts to devise a *sui generis* intellectual property right in noncopyrightable collections of data that would suitably balance public and private interests. It also emphasizes the need to reconcile legal protection of databases with fundamental constitutional mandates concerning free speech and the progress of science. The article concludes with a warning that overly protective initiatives could compromise the research-based institutions that currently ensure the technological predominance of U.S. industry in the global marketplace.

TABLE OF CONTENTS

| | | |
|-----|--|-----|
| I. | COMMODIFICATION OF DATA IN THE NETWORKED ENVIRONMENT: THE BIGGER PICTURE | 796 |
| II. | POTENTIAL IMPACT OF THE DATABASE PROTECTION LAWS ON SCIENCE AND TECHNOLOGY | 799 |

© 1999 J.H. Reichman and Paul F. Uhlir.

† J.H. Reichman, Professor of Law, Vanderbilt University School of Law and Paul F. Uhlir, National Research Council, Washington, DC. An early version of this paper was presented to the Symposium on *The Changing Character, Use, and Protection of Intellectual Property*, German-American Academic Council in Cooperation with the U.S. National Academy of Sciences and the Max Planck Institute for Foreign and International Patent, Copyright and Competition Law, Washington, D.C., Dec. 3-4, 1998, and to the Conference on *Law in the Information Society*, Istituto per la Documentazione Giuridica, Comitato Nazionale della Ricerca, Florence, Italy, Dec. 2-5, 1998. The views expressed in this article are the authors' and not necessarily those of the National Academy of Sciences or the National Research Council.

| | | |
|------|--|-----|
| A. | The User-Friendly Rules of Copyright Law | 800 |
| B. | Unbalanced Rules of the <i>Sui Generis</i> Model | 802 |
| 1. | <i>Basic Substantive Principles</i> | 803 |
| 2. | <i>The Resulting Legal and Practical Constraints</i> | 806 |
| C. | Long-term Implications of the <i>Sui Generis</i> Model | 812 |
| 1. | <i>Reversing the Transparency Movement</i> | 813 |
| 2. | <i>Transaction Costs Unlimited</i> | 814 |
| 3. | <i>Endless Monopolies and Diminished Access to Government Data</i> | 816 |
| 4. | <i>Gaming the Cooperative Ethos</i> | 819 |
| III. | RECENT DEVELOPMENTS: THE QUEST FOR AN APPROPRIATE UNFAIR COMPETITION APPROACH | 821 |
| A. | The Administration's Position | 822 |
| B. | A Negotiated Discussion Draft in the Senate | 823 |
| 1. | <i>Clarifying the Demands on Scientific and Technical Users</i> | 824 |
| 2. | <i>Compromise Proposals</i> | 825 |
| C. | Uncertain Future of the Database Protection Law | 829 |
| IV. | PRESERVING THE CONSTITUTIONAL BALANCE OF INTERESTS IN THE NETWORKED ENVIRONMENT | 832 |
| A. | The Competitive Ethos Under Attack | 832 |
| B. | The Constitutional Dilemma | 833 |
| C. | Erring on the Side of Caution | 836 |

INTRODUCTION

The convergence of digital computing and telecommunications technologies has greatly expanded the already bright economic prospects for information goods of all kinds, but it has also unsettled the legal architecture on which the free market economies have previously been grounded.¹ Information products behave differently from the tangible, physical products of the Industrial Revolution;² and the legal paradigms that we have applied to balance incentives to create against both public good uses of

1. See, e.g., Julie E. Cohen, *Lochner in Cyberspace: The New Economic Orthodoxy of "Rights Management,"* 97 MICH. L. REV. 462 (1998); Trotter Hardy, *Property (and Copyright) in Cyberspace*, 1996 U. CHI. LEGAL F. 217 (1996); Robert P. Merges, *The End of Friction? Property Rights and Contract in the "Newtonian" World of On-line Commerce*, 12 BERKELEY TECH. L.J. 115 (1997); Henry H. Perritt, Jr., *Property and Innovation in the Global Information Infrastructure*, 1996 U. CHI. LEGAL F. 261 (1996).

2. See, e.g., Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources to Invention*, in THE RATE AND DIRECTION OF INVENTIVE ACTIVITY 609, 616 (National Bureau of Economic Research ed., 1962) (stressing that optimal utilization occurs when information is free, while optimal information production occurs only when producers expect to appropriate the economic value of their investments); see also Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CALIF. L. REV. 479, 591-608 (1995) (describing the economics of networks as "still under construction").

information and the discipline of free competition are stretched past the breaking point.³ We are thus challenged to rethink how best to structure competition for information goods in the emerging, worldwide information economy.⁴

The technological convergence that creates promising new markets for information goods also opens new opportunities for scientific and educational uses of data and information. However, a powerful movement to commodify data and information previously treated as a public good—that is, as an inexhaustible, indivisible, and ubiquitous component of the public domain⁵—could limit the ability of the scientific, technical, and educational communities to capitalize on such opportunities. The momentum generated by that movement would eventually have faced these communities with serious challenges even in the absence of a new intellectual property right in collections of data. The adoption of a strong property right in noncopyrightable collections of data by the European Union⁶—in a haphazard manner, with little serious economic or empirical investigation⁷—thus precipitated a crisis that was already well under way.

3. See generally, J.H. Reichman, *Legal Hybrids Between the Patent and Copyright Paradigms*, 94 COLUM. L. REV. 2432 (1994) [hereinafter Reichman, *Legal Hybrids*]; J.H. Reichman, *Charting the Collapse of the Patent-Copyright Dichotomy: Premises for a Restructured International Intellectual Property System*, 13 CARDOZO ARTS & ENT. L. J. 475 (1995) [hereinafter Reichman, *Charting*]; Pamela Samuelson, Randall Davis, Mitchell D. Kapor & J.H. Reichman, *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308 (1994) [hereinafter Samuelson et al.].

4. See, e.g., Charles R. McManis, *Taking TRIPS on the Information Superhighway: International Intellectual Property Protection and Emerging Computer Technology*, 41 VILL. L. REV. 207 (1996).

5. See ROBERT COOTER & THOMAS ULEN, *LAW AND ECONOMICS* 40-41 (2nd ed. 1997) (noting that public goods are both non-excludable and non-rivalrous).

6. See Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases, 1996 O.J. (L 77) 20 [hereinafter E.U. Directive].

7. See J.H. Reichman & Pamela Samuelson, *Intellectual Property Rights in Data?*, 50 VAND. L. REV. 51, 72-95 (1997) (tracing legislative history of the E.U. Directive). For different perspectives, see, for example, Robert C. Denicola, *Copyright in Collections of Fact: A Theory for the Protection of Nonfiction Literary Works*, 81 COLUM. L. REV. 516 (1981) and Jane C. Ginsburg, *Copyright, Common Law, and Sui Generis Protection of Databases in the United States and Abroad*, 66 U. CIN. L. REV. 151 (1997); G.M. Hunsucker, *The European Database Directive: Regional Stepping Stone to an International Model?*, 7 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 697 (1997). See also Wendy J. Gordon, *On Owning Information: Intellectual Property and the Restitutory Impulse*, 78 VA. L. REV. 149 (1992).

This article explores the potentially adverse impact that the emerging legal infrastructure could have on scientific, technical and educational users of factual data and information (as well as on other sectors of the information economy) unless suitable adjustments are made. Parts I and II explain how efforts to accommodate the networked environment to the publishers' and database makers' fears of market failure will impose a daunting array of legal and contractual restraints on the ability of scientists and engineers to access factual data and information in the near future. Part III examines the most recent efforts to devise a *sui generis* intellectual property right in noncopyrightable collections of data that would suitably balance public and private interests. Part IV emphasizes the need to reconcile legal protection of databases with fundamental constitutional mandates concerning free speech and the progress of science. It ends with a warning that overly protective initiatives could compromise the research-based institutions that currently ensure the technological predominance of U.S. industry in the global marketplace.

I. COMMODIFICATION OF DATA IN THE NETWORKED ENVIRONMENT: THE BIGGER PICTURE

Digital telecommunications networks enable publishers to control the uses of information goods directly by contract, without relying on state action to avoid market failure,⁸ for the first time since the advent of the Guttenberg printing press. In effect, online delivery has "restored the power of the two-party deal" with regard to information goods and diminished the dependence of publishers on artificial legal fences that copyright laws and other related rights supplied in the print environment.⁹

Efforts to accommodate the pre-existing legal landscape to the new technologies are proceeding along several different fronts. For example, because the new technologies empower publishers to fence off information goods by means of encryption devices and other technical protection

8. See, e.g., PAUL GOLDSTEIN, *COPYRIGHT'S HIGHWAY: FROM GUTTENBERG TO THE CELESTIAL JUKEBOX* 27 (1994) (stating that "[c]opyright was technology's child from the start [because] [t]here was no need for copyright before the printing press"); Wendy J. Gordon, *Asymmetric Market Failure and Prisoner's Dilemma in Intellectual Property*, 17 U. DAYTON L. REV. 853 (1992) (relating economic justification of intellectual property rights to the problem of market failure).

9. See J. H. Reichman & Jonathan A. Franklin, *Privately Legislated Intellectual Property Rights: Reconciling Freedom of Contract with Public Good Uses of Information*, 147 U. PA. L. REV. 875, 897-99 (1999) (discussing "The Restored Power of the Two-Party Deal").

measures,¹⁰ Congress has been persuaded to pass new laws making it a civil or criminal offense to disarm or tamper with these devices.¹¹ Would-be users must increasingly gain access to information goods via an electronic gateway where they are obliged to identify themselves and acknowledge the rights of the gatekeeper to the information goods, as, for example, expressed in copyright management information systems.¹² The new laws that defend the owner's encryption devices also forbid users from tampering with their intellectual property identity tags.¹³

At the same time, the National Commissioners on Uniform State Laws have proposed a model contract law to govern computerized information transactions that all state legislatures would eventually adopt. These proposals, until recently embodied in a draft Article 2B of the Uniform Commercial Code ("U.C.C."),¹⁴ would validate the publishers' standard form, non-negotiable contracts to which would-be users must assent in order to cross the electronic threshold and gain access to information de-

10. See, e.g., Julie E. Cohen, *A Right to Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace*, 28 CONN. L. REV. 981, 983-89 (1996) (discussing technologies that copyright owners may utilize to monitor and control access to information); see also DanThu Thi Phan, Note, *Will Fair Use Function on the Internet?*, 98 COLUM. L. REV. 169, 192, 192 n.167 (1998) (defining and discussing "digital watermarks"). See generally NATIONAL RESEARCH COUNCIL, *CRYPTOGRAPHY'S ROLE IN SECURING THE INFORMATION SOCIETY* (Kenneth Dam & Herbert Lin eds., 1996); Mark Stefik, *Shifting the Possible: How Trusted Systems and Digital Property Rights Challenge Us to Rethink Digital Publishing*, 12 BERKELEY TECH. L. J. 137 (1997).

11. Digital Millennium Copyright Act, Pub. L. No. 105-304, §§ 103, 403, 112 Stat. 2860, 2863, 2889 (1998) (codified at 17 U.S.C. § 1201).

12. See Reichman & Franklin, *supra* note 9, at 897-98 (stating that this "gatekeeping function is reinforced by encryption devices, digital watermarking, and other self-help technical measures that permit information providers contractually to impose their own terms and conditions on access to information goods stored at any given network site and on the uses to which end-users can put the information they access").

13. See Digital Millennium Copyright Act, Pub. L. No. 105-304, § 103, 112 Stat. 2860, 2863 (1998) (codified at 17 U.S.C. § 1201).

14. See U.C.C. Article 2B—Licenses (Feb. 1999 Draft) (attempting to provide a common legal framework for transactions in digital information and software licenses). In April 1999, the American Law Institute, an original sponsor of this proposal, withdrew its support, and the National Commissioners announced their intention to pursue the project in the form of a model law governing computerized information transactions rather than as an amendment to the U.C.C. See NCCUSL & ALI, *NCCUSL to Promulgate Freestanding Uniform Computer Transactions Act: ALI and NCCUSL Announce that Legal Rules for Computer Information Will Not be Part of U.C.C.* (visited Apr. 23, 1999) <<http://www.nccusl.org/pressrel/2brel.html>>. For convenience, citations herein continue to refer to draft Article 2B, February 1999, the latest available version at the time of writing.

livered online.¹⁵ As matters stand, and despite mounting criticism from intellectual property scholars, the proposed model law would validate even "click on" or shrinkwrap licenses that ignored or attempted to override public interest exceptions that favored users or competitors, including the technical and scientific communities, under the pre-existing legal infrastructure.¹⁶ For example, such contracts could override the right to make non-infringing uses of copyrighted works or the right to reverse-engineer subpatentable innovation,¹⁷ and they could require payment for uses that courts had previously deemed fair uses under the federal copyright law.¹⁸

A third line of attack is to devise new intellectual property rights that, among other things, would serve to reduce potential tensions between state contract laws and the federal intellectual property system. As we shall see, copyright law expressly permits many uses of copyrighted works that publishers would like to restrict by means of online licenses.¹⁹ The validity of such contracts may be questioned on the grounds that they disrupt the federal intellectual property system (preemption arguments) or overstep the constitutional guarantees of free thought and expression.²⁰

However, if Congress enacted a hybrid ("*sui generis*") intellectual property right to protect the contents of databases, like that adopted by the European Union, it would give legislative approval to forms of protection that were previously unknown or questionable under traditional intellec-

15. Raymond T. Nimmer, the Reporter for Article 2B, justifies this approach. See Raymond T. Nimmer, *Breaking Barriers: The Relation Between Contract and Intellectual Property Law*, 13 BERKELEY TECH. L. J. 827 (1998).

16. See, e.g., Julie E. Cohen, *Copyright and the Jurisprudence of Self-Help*, 13 BERKELEY TECH. L.J. 1089 (1998); Jessica Litman, *The Tales that Article 2B Tells*, 13 BERKELEY TECH. L.J. 931 (1998); Charles R. McManis, *The Privatization (or "Shrink-Wrapping") of American Copyright Law*, 87 CALIF. L. REV. 173 (1999).

17. See, e.g., Reichman & Franklin, *supra* note 9, at 939-43, 947-51; see also Rochelle Cooper Dreyfuss, *Do You Want to Know a Trade Secret? How Article 2B Will Make Licensing Trade Secrets Easier (But Innovation More Difficult)*, 87 CALIF. L. REV. 191 (1999); Mark A. Lemley, *Beyond Preemption: The Law and Policy of Intellectual Property Licensing*, 87 CALIF. L. REV. 111 (1999).

18. See McManis, *supra* note 16, at 173 (addressing the capacity of Article 2B to alter the existing balance embodied in copyright law); Reichman & Franklin, *supra* note 9, at 943-47.

19. See 17 U.S.C. §§ 102(b) (1994) (ideas not protectable), 107 (1994) (fair use), 108-121 (1994) (other exceptions and limitations).

20. See 17 U.S.C. § 301 (1994); see also Dennis S. Karjala, *Federal Preemption of Shrinkwrap and On-line Licenses*, 22 U. DAYTON L. REV. 511 (1997); David A. Rice, *Public Goods, Private Contract and Public Policy: Federal Preemption of Software License Prohibitions Against Reverse Engineering*, 53 U. PITT. L. REV. 543 (1992).

tual property law.²¹ The creation of new intellectual property rights in collections of data would thus make it harder to resist arguments that publishers who subject online delivery of databases to technical protection measures and to contracts of adhesion that limit previously legal uses had violated fundamental public policies derived from the copyright laws. In other words, the database protection laws seem to permit acts (and foster policies) that overtly contradict or override the limits previously established by copyright and other traditional legal models.²²

Taken together, the ability of publishers to combine technical protection measures with tailor-made contract laws and hybrid intellectual property rights is supposed to stimulate investment in online commerce and to foster overall economic development.²³ Critics fear, however, that the cumulative effect of these separate but well-coordinated legal initiatives will be to balkanize the information economy and to unduly restrict the use of unbundled information as raw materials of science and technology or as inputs into the production of value-adding or second generation information goods.²⁴

II. POTENTIAL IMPACT OF THE DATABASE PROTECTION LAWS ON SCIENCE AND TECHNOLOGY

Let us suppose that a scientist or engineer lawfully obtained a printed copy of a chemical handbook or of a scientific article, with appended data, that was published in a peer-reviewed journal. These works currently at-

21. See E.U. Directive, *supra* note 6, arts. 7-10. For U.S. proposals spawned by the E. U. Directive, see Collections of Information Antipiracy Act, H.R. 354, 106th Cong. (1999); Collections of Information Antipiracy Act, H.R. 2652, 105th Cong. (1998); Database Investment and Intellectual Property Antipiracy Act of 1996, H.R. 3531, 104th Cong. (1996).

22. See Reichman & Samuelson, *supra* note 7, at 84-95, 103-110.

23. See, e.g., Nimmer, *supra* note 15; Hardy, *supra* note 1.

24. See *supra* notes 16-17; Reichman & Franklin, *supra* note 9, at 881 (stating that, to "ignore such discriminations as these is to risk watching model laws, adopted to govern the virtual marketplace for information goods, foster conditions that actually decrease innovation, discourage competition, and stifle the traditional marketplace of ideas"); see also G.E. Evans & B.F. Fitzgerald, *Information Transactions Under U.C.C. Article 2B: The Ascendancy of Freedom of Contract in the Digital Millenium?*, 21 U. NEW S. WALES L.J. 404 (1998) (arguing for the government's need to protect society given the recent shift of market power in the information economy).

tract copyright protection, and we shall assume that they meet the eligibility criteria of that body of law.²⁵

A. The User-Friendly Rules of Copyright Law

The rules of copyright law constitute a balanced regime of public and private interests. In retrospect, we are struck by the friendly treatment this body of law gives to users and competitors alike, notwithstanding the powerful bundle of exclusive rights it vests in authors and artists in order to stimulate the production and dissemination of creative works.²⁶

For example, any scientist or engineer who lawfully obtained the book or article mentioned above could immediately re-use all the data and all the ideas disclosed in them because copyright law does not protect ideas or data,²⁷ nor does it protect against use of expression as such, but only against certain specified uses.²⁸ Indeed, another scientist or engineer could independently rewrite his or her own version of the same article and disseminate it because copyright law allows independent creation, and all the unprotected data are spread out before the second comer's eyes.²⁹

A second scientist or engineer who needed to duplicate even the first author's creative selection and arrangement of data (if any) for non-profit research purposes could normally fall back upon the "fair use" provisions of current law.³⁰ A later researcher could also produce a follow-on article

25. See 17 U.S.C. §§ 101 (1994) (defining literary works), 102(a) (1994) (requiring original works of authorship), 103(b) (1994) (protection limited to original and expressive material added by author to a compilation); *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991) (limiting copyright protection of factual compilation to creative elements of selection, arrangement, and coordination).

26. See 17 U.S.C. § 106 (1994). See generally L. RAY PATTERSON & STANLEY W. LINDBERG, *THE NATURE OF COPYRIGHT: A LAW OF USERS' RIGHTS* 47-56, 191-224 (1991); JOEL SHELTON LAWRENCE & BERNARD TIMBERG, *FAIR USE AND FREE INQUIRY: COPYRIGHT LAW AND THE NEW MEDIA* (2d ed. 1989).

27. See *supra* note 25; *Harper & Row v. Nation Enterprises*, 471 U.S. 539 (1985) (stressing First Amendment interest in unrestricted availability of facts).

28. See 17 U.S.C. §§ 106, 106A (1994); *Baker v. Selden*, 101 U.S. 99 (1879); Ralph S. Brown, Jr., *Eligibility for Copyright Protection: A Search for Principled Standards*, 70 MINN. L. REV. 579, 588-89 (1985) (noting that the Copyright Act, unlike the Patent Act, does not confer any exclusive right to use the protected work, which helps explain why copyrights are so casually granted).

29. See, e.g., PAUL GOLDSTEIN, *COPYRIGHT* § 7.2.2 (2d ed. 1998) (noting that "conveying evidence" of independent creation constitutes a perfect defense to an infringement action).

30. See 17 U.S.C. § 107 (1994) (codifying fair use provisions); *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 564, 574-594 (1994) (stressing desirability of promoting transformative, preambular uses under the fair use provision of § 107); WILLIAM F.

or book that borrowed the originator's unprotected factual information and data, but not his or her stylistic expression.³¹ To be sure, the norms of world copyright law (but not necessarily U.S. law) favor attribution in such a case, as do the ethics of science.³² But plagiarism is not the same as copyright infringement; the reuse of facts and data is clearly permitted in copyright law; and another author's popularized version of a prior researcher's factual findings remains perfectly legal.³³

Most important, later scientists could combine the published data and factual information with other data and information into a multiple or complex interdisciplinary database without permission or additional payment to the originators.³⁴ This follows in part because only ineligibility matter is at issue and in part because copyright law does not prohibit *use* as such, but only certain uses, such as reproduction or adaptation of protected expression, and it is also buttressed by the doctrine of fair use.³⁵

Even if scientists, engineers, or educators made classroom use of the protected expression for nonprofit purposes, these uses might well be fair

PATRY, *THE FAIR USE PRIVILEGE IN COPYRIGHT LAW* 178-84, 416-17 (1985). However, the wholesale duplication of a creative selection and arrangement for commercial purposes is not permitted. *See, e.g.*, *CCC Info. Servs., Inc. v. Maclean Hunter Mkt. Reports, Inc.*, 44 F.3d 61 (2d Cir. 1994); *Reichman & Samuelson, supra* note 7, at 63 (citing authorities). *But see* *Warren Publishing Inc. v. Microdos Data Corp.*, 115 F.3d 1509 (11th Cir. 1995) (en banc) (allowing massive extraction and re-use of commercial compilation for competitive purposes and finding no eligibility for copyright protection).

31. *See* 17 U.S.C. §§ 102(b), 103(b) (1994); *Harper & Row v. Nation Enters.*, 471 U.S. 539 (1985); *Key Publications, Inc. v. Chinatown Today Publ'g Enters., Inc.*, 945 F.2d 509, 514 (2d Cir. 1991) (stressing "thin" copyright protection doctrine of *Feist*). *See also* *Reichman & Samuelson, supra* note 7, at 63 (citing authorities).

32. *See* *Berne Convention for the Protection of Literary and Artistic Works*, Sept. 9, 1886, as last revised at Paris, July 24, 1971, 828 U.N.T.S. 4, art. 6bis (obliging member states to respect moral rights of authors); *Berne Convention Implementation Act of 1988*, Pub. L. No. 100-568, 102 Stat. 2853 (1988) (implementing this obligation only indirectly).

33. *See* *Hoehling v. Universal City Studios, Inc.*, 618 F.2d 972 (2d Cir. 1980); *supra* notes 31-32 and accompanying text. However, if a second comer were to pass off his article as that of the first scientist, grounds for invoking relief in unfair competition law might also exist. *See* *Lanham Act § 43(a)*, 15 U.S.C. § 1125 (1998).

34. For the importance of this practice, in conjunction with the sharing ethos of science, see J. H. Reichman, *Why Science is Concerned About an Intellectual Property Right in Databases*, in *AAAS SCIENCE AND TECHNOLOGY POLICY YEARBOOK 1998* (Albert H. Teich et al. eds., 1998), at 291, 301; *see also* *International Council for Science (ICSU), Position Paper on Access to Databases*, paper presented to the World Intellectual Property Organization (Sept. 1997) (unpublished manuscript, on file with authors).

35. *See supra* notes 26-31 and accompanying text.

or privileged uses under U.S. copyright law³⁶ and would possibly become subject to compulsory licensing under E.U. copyright laws.³⁷ Finally, having once purchased the book or the article, a scientist or engineer could sell it, lend it, or give it to others (first sale doctrine),³⁸ borrow it from a library,³⁹ use it as often as he or she liked for virtually any purpose, and make photocopies of it for scientific purposes under the fair use doctrine of U.S. law⁴⁰ or the private use doctrine of E.U. law.⁴¹

B. Unbalanced Rules of the *Sui Generis* Model

Now, let us suppose that the contents of the same chemical handbook or of the aforementioned scientific article were disseminated online and surrounded by technical fences as previously described. Suppose further that the contents of the book or article were protected by laws implementing the E.U.'s *sui generis* exclusive property right in noncopyrightable collections of data or by the U.S. version of that right, as set out in H.R. 2652, the Collections of Information Antipiracy Act (March 1998).⁴² The House of Representatives adopted H.R. 2652, and the Subcommittee on Courts and Intellectual Property then attached it to the Digital Millennium

36. See 17 U.S.C. § 107, 110(1) (1994); *but see* Princeton Univ. Press v. Michigan Document Servs., Inc., 99 F.3d 1381 (6th Cir. 1996) (en banc), *cert. denied*, 520 U.S. 1156 (1997) (limits on copies for classroom use).

37. See, e.g., Berne Convention, *supra* note 32, arts. 9(2) (limits on reproduction right), 10(2) (use for teaching purposes when consistent with fair practice); GUY TRITTON, INTELLECTUAL PROPERTY IN EUROPE 191 (1996). See also Lucie Guibault, General Report to the ALAI Annual Meeting (1998) (unpublished manuscript, on file with authors) (discussing exceptions and limitations in European copyright law).

38. See 17 U.S.C. § 109(a) (1994) (first sale doctrine).

39. See *id.*; see also 17 U.S.C. § 108 (1994) (reproduction by libraries and archives).

40. See 17 U.S.C. § 107 (1994); *supra* note 30. However, scientists and engineers working at for-profit institutions have lesser photocopying privileges, at least when secondary markets for photocopies and reprints are operational. See American Geophysical Union v. Texaco, Inc., 60 F.3d 913, 926-931 (2d. Cir. 1994); PATRY, *supra* note 30, at 190-94.

41. See European Commission Green Paper on Copyright and Related Rights in the Information Society, *reprinted in* 43 J. COPYRIGHT SOC'Y USA 50, 91 (1995) ("Most member states have introduced special legal arrangements for ... private copying..."). *But see* William R. Cornish, *Copyright in Scientific Works (Scientific Communications, Computer Software, Data Banks): An Introduction*, in EUROPEAN RESEARCH STRUCTURES—CHANGES AND CHALLENGES: THE ROLE AND FUNCTION OF INTELLECTUAL PROPERTY RIGHTS 50 (Max Planck Gesellschaft ed., 1994) (despite case for a measure of free reprography for purposes of academic research, "academic institutions are regarded as relatively soft targets by publishing interests [in U.K.], which ... [have been] inserting initial wedges.").

42. See E.U. Directive, *supra* note 6; H.R. 2652, 105th Cong. (1998).

Copyright Act, which became H.R. 2281, as sent to the Senate.⁴³ The database portion was dropped prior to Congressional enactment of that bill, however, and it was reintroduced with some modifications as H.R. 354 in January 1999.⁴⁴

1. *Basic Substantive Principles*

The *sui generis* provisions of the E.U. Directive⁴⁵ protect the contents of any noncopyrightable database that is the product of substantial investment against *extraction or reutilization* of the whole or of any substantial part (evaluated quantitatively or qualitatively).⁴⁶ Hence, this law could protect the noncopyrightable data appended to the hypothetical article in question or collected in the handbook, which the publishers might eventually disseminate online, with or without an accompanying print version.⁴⁷

Such protection lasts as long as new investments are made in updates or maintenance; hence perpetual protection of dynamic databases becomes a likely result, despite a nominal fifteen-year term.⁴⁸ There are no exceptions for "reutilization" by scientific and educational bodies, and there are no mandatory exceptions for "extraction" for scientific and educational purposes (although states may adopt this exception for noncommercial purposes).⁴⁹ The member states implementing the Directive must permit

43. See H.R. 2281, 105th Cong. (1998) (including the database protection bill as Title V).

44. See H.R. 354, 106th Cong. (1999).

45. See E.U. Directive, *supra* note 6, arts. 7-10. Besides a list of sixty "Recitals" or premises that underlie the legislation and a small set of definitional articles that apply across the board (arts. 1-2), the E.U. Directive also harmonizes the treatment of copyrightable databases in the member states' domestic laws. See *id.* arts. 3-6; Reichman & Samuelson, *supra* note 7, at 76-79. There is a final group of "common provisions" that apply to both copyrightable and noncopyrightable databases (arts. 12-16). While the provisions harmonizing the protection of copyrightable databases approximate the rules in the U.S., see *supra* notes 26-41 and accompanying text, they lie beyond the scope of this article.

46. See E.U. Directive, *supra* note 6, art. 7(1).

47. For the very broad definition of databases, see E.U. Directive, *supra* note 6, art. 1. For example, Reed Elsevier, Inc. has been buying up scientific journals and has recently begun to deliver scientific communications online.

48. See E.U. Directive, *supra* note 6, art. 10; Reichman & Samuelson, *supra* note 7, at 85-86.

49. See E.U. Directive, *supra* note 6, arts. 9, 9(b) (authorizing member states to allow extractions (but not reutilization) "for the purposes of illustration for teaching or scientific research, so long as the source is indicated and to the extent justified by the non-commercial purpose to be achieved").

extraction or use of an insubstantial part of a protected database.⁵⁰ However, the risks of invoking even this exception are high, because a would-be user has no way of knowing in advance whether a court will later find that the amount used was in fact qualitatively or quantitatively insubstantial.⁵¹

U.S. bill H.R. 2652, later H.R. 2281 as part of the Digital Millennium Copyright Act, used different language to accomplish essentially the same result. It protected against *use or extraction* in commerce of all or a substantial part of a protected collection of information that is the product of substantial investment if such use or extraction would "cause harm to the actual or potential market" for a product or service that incorporated the collection.⁵² For this purpose, the term "collection of information" was very broadly defined,⁵³ and during face-to-face negotiation in the Senate (in which the authors of this article participated directly),⁵⁴ the publishers claimed that a single lost sale would fit within this standard of harm to the market. Any substantial new investment in updates or maintenance would prolong protection beyond fifteen years, with no limit to the number of renewals.⁵⁵ The bill initially recognized no exceptions for science and education as such, but, at the last minute, a provision tacked onto H.R. 2281 held scientists and educators liable only for harm to "actual mar-

50. See E.U. Directive, *supra* note 6, art. 8(1). Member states may allow a broader exception for "extraction for private purposes of the contents of a *non-electronic database*." *Id.* art. 9(a) (emphasis added).

51. See Reichman & Samuelson, *supra* note 7, at 90-91. See also *id.* at 87-95 (finding that the scope of protection under the E.U. Directive exceeds that of copyright law because 1) no idea-expression distinction is observed and no evolving public domain is generated, 2) the equivalent of a "derivative work" right in dynamic databases is not limited to new matter, and 3) the public interest exceptions are very narrow).

52. H.R. 2652, 105th Cong. § 1202 (1998); H.R. 2281, 105th Cong. § 1302 (1998).

53. See H.R. 2281, 105th Cong. § 1301(1) (1998) (defining "collection of information" to mean "information that has been collected and ... organized for the purpose of bringing discrete items of information together in one place or through one source so that users may access them"); *id.* § 1301(2) (defining "information" to mean "facts, data, works of authorship, or any other intangible material capable of being collected and organized in a systematic way"). See also *id.* § 1301(3) (defining "potential market" to mean "any market that a person claiming protection ... has current and demonstrable plans to exploit or that is commonly exploited by persons offering similar products or services incorporating collections of information").

54. See *infra* text accompanying notes 143-47.

55. See H.R. 2281, 105th Cong. § 1308(c) (1998). The interpretation in the text was confirmed by the position that the publishers took during face-to-face negotiations in the Senate. See *infra* text accompanying notes 143-47.

kets,” and not for harm to “potential markets” for their nonprofit uses of protected information.⁵⁶

The situation was further complicated in January, 1999, when Chairman Howard Coble of the House Committee on the Judiciary’s subcommittee dealing with intellectual property rights introduced a new version of the database protection bill, H.R. 354,⁵⁷ which modified the previous bill in at least two important respects. First, a serious effort was made to limit the term of protection to fifteen years, with little or no possibility of extension even in dynamic databases that are continuously updated.⁵⁸ Second, a new provision established an exemption for “additional, reasonable uses” by educational, scientific and research organizations,⁵⁹ which was loosely based on the “fair use” provisions of copyright law. However, this ambiguous provision would limit the proposed exception to “an individual act of use or extraction of information done for” specified purposes,⁶⁰ which apparently placed the burden of proof on the otherwise infringing researcher.⁶¹

Because one can read the new exception for scientific and educational uses set out in H.R. 354 broadly or narrowly, depending on how one interprets its latent ambiguities, it is instructive to assess the likely impact of the proposed legislation on science and technology as it stood at the end of 1998. We can then factor the proposed amendments into the analysis and compare them to certain promising proposals that emerged from face-to-face negotiations between stakeholders, held in the Senate under Senator

56. See H.R. 2281, 105th Cong. § 1303(d) (1998). An article attacking the proposed legislation that appeared in *Science* magazine focused particular attention on the harm to science from a “potential market” test. See William Gardner & Joseph Rosenbaum, *Database Protection and Access to Information*, 281 *SCIENCE* 786-87 (1998).

57. H.R. 354, 106th Cong. (1999). See also *Hearings on H.R. 354, the “Collections of Information Antipiracy Act” before the House Subcomm. on Courts and Intellectual Property of the House Comm. of the Judiciary*, 106th Cong. (1999) [hereinafter *Hearings*] (statement of the Honorable Howard Coble, Chair of the Subcommittee on Courts and Intellectual Property).

58. See H.R. 354, 106th Cong. § 1408(c) (1999) (limiting the term of protection to 15 years).

59. *Id.* § 1403(a)(2) (listing “Additional Reasonable Uses”).

60. *Id.* § 1403(a)(2)(A) (allowing individual act of use or extraction “for the purpose of illustration, explanation, example, comment, criticism, teaching, research, or analysis, in an amount appropriate and customary for that purpose ... if [such an act] is reasonable under the circumstances”) Criteria for determining reasonable use, and the limits thereon, are also set out. See *id.* § 1403(a)(2)(A)(i)-(iv).

61. See *infra* text accompanying notes 95-99.

Hatch's auspices in late summer of 1998.⁶² Accordingly, if Congress had adopted H.R. 2281 at the end of 1998, and that law had subsequently been applied to online delivery of the data contained in the book or article that were previously discussed in connection with the workings of copyright law,⁶³ the following results would have been likely to occur.

2. *The Resulting Legal and Practical Constraints*

In principle, a second scientist or engineer could not make any uses of the information or data that were not permitted by the form-contract site licenses that regulated access to the online database from which they were extracted.⁶⁴ The site license could charge one price for accessing or consulting the database, a second price for downloading it, and a third price for using it or reusing it in other contexts.⁶⁵

62. See *infra* text accompanying notes 140-65.

63. See *supra* text accompanying notes 26-41.

64. While this constraint could occur under existing law, the validity of such standard form contracts remains in doubt at the present time, with some courts upholding them and other courts invalidating them either on contracts grounds or under the doctrine of preemption. See Reichman & Franklin, *supra* note 9, at 876 n.1 (citing cases). The adoption of a model law governing computerized information transactions (like the previously proposed Article 2B of the U.C.C.), see *supra* note 14, would validate virtually all such contracts. See *id.* at 899-914 (criticizing this approach and proposing a new doctrine of "public interest unconscionability" to allow courts to reconcile freedom of contract with public-good uses of information). However, most opportunities to challenge the validity of such contracts as applied to either copyrightable or noncopyrightable databases on existing grounds would vanish if Congress adopted a database law along the lines of H.R. 2281, unless some other countervailing doctrine, such as the proposed "public interest unconscionability doctrine" became available. See Reichman & Franklin, *supra* note 9, at 947-51 ("Contracts Restricting the Use of Noncopyrightable Collections of Data").

If, instead, H.R. 354 were adopted, or a variant thereof that included a substantial exception for traditional scientific purposes, see *infra* text accompanying notes 148-52, the opportunities to challenge the validity of such contracts would depend on 1) the extent to which the database law itself restricted contractual overrides (none yet proposed, and all vigorously resisted by the publishers), and 2) the availability of ancillary doctrines in contracts law, such as the proposed "public interest unconscionability" doctrine. See *infra* text accompanying notes 163-64 (discussing the doctrine of misuse); see generally Reichman & Franklin, *supra* note 9, at 929-38 ("validating non-negotiable terms that respect the balance of public and private interests").

65. Cf. Reichman & Samuelson, *supra* note 7, at 117-24 ("Retarding the Progress of Science").

[T]he electronic publisher's growing capacity to charge for each and every use of online data (or at least for every "hit" that accesses such data), and to track and monitor every user ... means that it becomes in-

Even though the second scientist or engineer normally would have paid to access the data and information (and they are not copyrightable by definition), he or she could not use them in ways not permitted by the terms and conditions of that site license, which, in turn, would now be supported by a duly enacted federal intellectual property law.⁶⁶ Absent some constitutional override, the second comer could not, therefore, independently generate a similar article or study based on the same material without permission, even though the relevant data were now revealed to the public.⁶⁷ Because the data no longer entered the public domain,⁶⁸ he or she would need to obtain a new grant or substitute funding to repeat the collection process, in which case scarce funds would have been used to duplicate the creation of knowledge already in existence. This, of course, contradicts the norms of science, which favor building on previous discoveries and the sharing of research results.⁶⁹

In many instances, the data will be based on one-time events that later scientists and engineers could not physically regenerate, in order to fall within the permitted acts of independent creation under the database protection laws.⁷⁰ Even when regeneration remained feasible, the cost in rela-

creasingly capable of serving ... as its own collection society, subject to no consent decrees ... and no external regulation.

Id. at 153. At a recent conference on database protection in Italy that Professor Reichman attended, an Italian professor related that the European Commission had charged him a high price per page to consult official texts concerning antitrust laws and regulations and about double that price per page to download the same information for research purposes. See Tito Ballarino, Remarks at the University of Pavia Conference on "Le banche dati (anche su Internet)," Pavia, Italy, Oct. 2-3, 1997; cf. Jessica Litman, *The Exclusive Right to Read*, 13 CARDOZO ARTS & ENT. L. J. 29 (1994).

66. See *supra* note 64.

67. See H.R. 2281, 105th Cong. § 1303(b) (1998) (declining to "restrict any person from independently gathering ... or using information gathered ... by another person through the investment of substantial monetary or other resources"); *id.* § 1305(e) (allowing unrestricted licensing agreements). See also H.R. 354, 106th Cong. §§ 1403(c), 1405(e) (1999) (the same in this respect, but attenuated in impact owing to fair-use-like provisions set out in § 1403(a)). For possible constitutional overrides, see *infra* text accompanying notes 182-97.

68. Cf. Reichman & Samuelson, *supra* note 7, at 84-90 ("Abolishing the Public Domain"). Under H.R. 354, however, data would enter the public domain after fifteen years. See *supra* note 58.

69. See NATIONAL RESEARCH COUNCIL, BITS OF POWER: ISSUES IN GLOBAL ACCESS TO SCIENTIFIC DATA 1, 132 (1997) (stressing importance of scientific norms that favor the sharing of data and the cumulative process of acquiring scientific knowledge) [hereinafter BITS OF POWER].

70. See *supra* note 67.

tion to the niche market of likely users would normally be so high that few second comers would willingly regenerate the data.⁷¹ Hence, sole-source providers are likely to remain a dominant feature of the database landscape, real competition will continue to be the exception, and the strong property rights given database proprietors would potentiate existing barriers to entry.⁷²

Later scientists and engineers could not combine data legitimately accessed from one commercial database with data extracted from other databases to make a complex new database for addressing hard problems without obtaining additional licenses and permissions. This remains, perhaps, the single most critical problem for scientific and technical research.⁷³ Despite reassurances to the contrary from leaders of the international publishing community to leaders of the scientific community at a recent meeting in Paris,⁷⁴ lawyers representing publishers at face-to-face negotiations held in the Senate late in 1998 continued to insist that this customary and traditional scientific practice would, in principle, violate their redistribution rights.⁷⁵ Another critical factor is that there would never be a sale that exhausted the publisher's rights, only a license, which the proposed model laws of computerized information transactions would make perpetual.⁷⁶

No one could combine "substantial" amounts of data or information into a more efficient follow-on product without a license; the licensor

71. See, e.g., BITS OF POWER, *supra* note 69, at 114-24.

72. See *id.*; Reichman & Samuelson, *supra* note 7, at 90-95 ("Establishing Legal Barriers to Entry"), 124-30 ("Impeding Competition in the Market for Value-Adding Products and Services"). But see Laura D'Andrea Tyson & Edward F. Sherry, Statutory Protection for Databases, in *Hearing on H.R. 2652 Before the Subcomm. on Courts and Intellectual Property of the House Comm. on the Judiciary*, 105th Cong. (1997) (testimony of Laura D'Andrea Tyson, Consultant, Reed-Elsevier, Inc.) (contesting the strength of barriers to entry in database industry). New studies by the National Research Council later in 1999 will attempt to cast further light on these issues.

73. See ICSU Position Paper, *supra* note 34; *infra* notes 110-14 and accompanying text.

74. Trip report by Ferris Webster, Chair, ICSU/CODATA Group on Access to Data and Information, International Council for Science (1998) (on file with authors).

75. Whether the new exceptions proposed in H.R. 354 would alter this result remains to be seen. See *supra* notes 58-60, and accompanying text. See also *infra* text accompanying notes 140-46.

76. See U.C.C. § 2B-502 (Feb. 1999 Draft) (allowing prohibition of any transfer of mass-market licensed goods); Karjala, *supra* note 20, at 538; Reichman & Franklin, *supra* note 9, at 965. Unless otherwise restrained, such licenses could override the limited duration clause that H.R. 354 finally introduced. See *supra* note 58.

would labor under no duty to grant such a license: and the sole-source provider would not want any competition from follow-on products.⁷⁷ This also suggests, however, that the price would not be set so high as to encourage independent creation of the same data, when otherwise feasible. If so, and potential producers of follow-on products tended to invest in other activities, it would further discourage competition and innovation.⁷⁸

So long as natural and artificial barriers to entry remained high, scientists and engineers must pay artificially high prices to access commercial databases in the absence of competition. The enactment of strong exclusive property rights (complemented by strengthened contractual rights if the proposed model law were also adopted⁷⁹) thus seems likely to reinforce the pervasive sole-source character of the marketplace and exert further upward pressure on prices.⁸⁰

Meanwhile, scientists and engineers who paid to access protected databases could not routinely lessen overall transaction costs by lending, borrowing, or transferring the data they extracted to others working on a common problem. This follows because there would never be a sale or transfer under some equivalent of the "first sale" doctrine of the copyright (and patent) laws,⁸¹ only a license that would logically restrict further transfers without any time limit.⁸² Scientists and engineers who continued to share data once acquired without obtaining permission and without paying additional fees for such heretofore traditional or customary uses would "harm the market" that the database proprietor presumably secured by dint of the proposed legislation.⁸³

The data would not enter the public domain for at least fifteen years, and possibly never, if the private party were to continue to invest in main-

77. See Reichman & Samuelson, *supra* note 7, at 124-30. Cf. *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991) (sole-source proprietor of telephone directory denied permission to another directory compiler who wished to combine the data in the former's directory with data from numerous others whose owners had given their consent).

78. See *supra* notes 71-72.

79. See *supra* notes 14-15 and accompanying text.

80. Cf. BITS OF POWER, *supra* note 69, at 121-23 (chronicling failed attempt to privatize Landsat data in the 1980s, when prices of Landsat images rose from about \$400 per image to \$4,400 per image, a price at which the joint venture "was able to attract some commercial and federal customers, but few academic or independent researchers").

81. See *supra* notes 38, 76.

82. See Reichman & Franklin, *supra* note 9, at 964-65.

83. See H.R. 2281, 105th Cong. § 1302 (1998).

tenance or updates of a dynamic database.⁸⁴ Even data that nominally entered the public domain at expiry of the fifteen year term could remain unavailable in practice if would-be users lacked means to identify and isolate those data within the larger mix of protected and unprotected data comprising a dynamic collection.⁸⁵ If such data were rendered technically identifiable, nothing would prevent the proprietor from using electronic fencing devices and standard form contracts to further preclude extraction even after the intellectual property right had expired.⁸⁶

Moreover, unless proper precautions are taken, there is considerable risk that data generated or funded by the U.S. government would become privatized in ways that unduly restricted access on onerous terms and conditions.⁸⁷ If this were allowed to happen, taxpayer-financed data would be sold back to science and education at monopoly prices, with the likelihood that additional state subsidies would be needed to defray the costs. In the European Union, where governments intend to commercialize publicly funded data, insufficient thought has been given to this problem in general and to the impact on science and technology in particular.⁸⁸

A common thread uniting all the foregoing observations is the lack of any limits on the power of providers who benefit from legal protection of databases to impose any licensing terms or conditions they wish on access to, and use of, their products. In principle, the database provider could override by contract even the few exceptions and limitations contained in the bill, including the public's right to use insubstantial parts of a database.⁸⁹

84. *See id.* § 1308(c).

85. *See* Jane C. Ginsburg, U.S. Initiatives to Protect Works of Low Authorship, paper presented to New York University Conference on "Intellectual Products: Novel Claims to Protection and Their Boundaries," Engelberg Center on Innovation Law and Policy, La Pietra, Italy (June 25-28, 1998) (unpublished manuscript, on file with authors) (arguing that publishers should identify the expired components of protected compilations).

86. *See* Reichman & Franklin, *supra* note 9, at 897-913, 947-51.

87. *See, e.g.,* *Hearings, supra* note 57 (statement of Andrew J. Pincus, General Counsel, U.S. Department of Commerce), at 13-20 (arguing for broad exemptions for government-funded data and warning about the "potential for "capture" of government-generated data) [hereinafter Statement of Pincus].

88. *See, e.g.,* Peter N. Weiss & Peter Backlund, *International Information Policy in Conflict: Open and Unrestricted Access versus Government Commercialization*, in *BORDERS IN CYBERSPACE* 300, 303 (Brian Kahin & Charles Nesson eds., 1997). *See also supra* note 65.

89. *See* H.R. 2281, 105th Cong. § 1303(a), 1305(e) (1998). *But cf.* E.U. Directive, *supra* note 6, art. 8(1). This E.U. privilege may not be overridden by contract, *see id.* art.

The net result, as Professors Reichman and Samuelson pointed out in an earlier article, is that, under the U.S. database proposals, as under the E.U. Directive,

the most borderline and suspect of all the objects of protection ever to enter the universe of intellectual property discourse—raw data, scientific or otherwise—paradoxically obtains the strongest scope of protection available from any intellectual property regime except, perhaps, for the classical patent paradigm itself.⁹⁰

When the provisions added to the latest bill, H.R. 354,⁹¹ are factored into the analysis, the end result is only slightly improved, at least in appearance if not in practice.

The first significant change mentioned above, which would more clearly inject protected data into the public domain after fifteen years,⁹² is of course a move in the right direction. However, the drafters still ignore the difficulties of identifying and accessing data whose term of protection had technically expired, an issue that was widely discussed last year.⁹³ The bill also ignores the power of database providers to override formal access to data that nominally entered the public domain by combining adhesion contracts with electronic fencing devices.⁹⁴

The second major change is a good faith effort to address some of the concerns of the scientific and educational communities by means of new, "fair-use-like" provisions.⁹⁵ However, these provisions are both ambiguous and too narrowly drawn.⁹⁶ By placing the burden of proof on scientists

15. United States publishers opposed the ban on contractual overrides of this provision in the E.U. Directive, and indicate that they intend to override it when permitted. See Mark Powell, *The European Union's Database Directive: An International Antidote to the Side Effects of Feist*, paper presented to the Fourth Annual Conference on "International Intellectual Property Law & Policy," Fordham University School of Law (Apr. 11-12, 1996) (unpublished manuscript, on file with authors).

90. Reichman & Samuelson, *supra* note 7 at 94.

91. H.R. 354, 106th Cong. (1999). See *supra* notes 58-61 and accompanying text.

92. See *id.* § 1408(c) (1999) (limiting the term of protection to 15 years).

93. See *supra* note 85 and accompanying text; Statement of Pincus, *supra* note 87, at 25-27 (questioning ability of users "to distinguish unprotected data entries from protected data entries" and fearing "*de facto* perpetual protection").

94. See H.R. 354, 106th Cong. § 1405(e) (1999) (allowing freedom of contract); Reichman & Franklin, *supra* note 9, at 899-914.

95. See H.R. 354, 106th Cong. § 1403(a) (1999).

96. See, e.g., *Hearings, supra* note 57 (statement of James G. Neal, Dean, University Libraries, Johns Hopkins University) ("[E]xemption for education and research ... remains far too narrow.") [hereinafter Statement of Neal]; *id.*, (testimony of Charles E.

and engineers, whose “individual acts” of “reasonable” use remain subject to scrutiny case by case,⁹⁷ they would continue to exert the chilling effect on research⁹⁸ that seems inherent in any “fair use” approach to a database law that does not otherwise provide the many other safeguards familiar from copyright law. Hence, as we explain below, a different kind of approach, one not strictly linked to the “fair use” concept, will be needed to ensure that a *sui generis* database regime does not harm customary and traditional scientific activities.⁹⁹

Even if a satisfactory legal formula to avoid harm to science and education were found, that formula would remain largely ineffective if database providers could simply override it by contract or, in the alternative, if the publishers could just charge more for access if they knew that the state would require them to charge less for extractions and reuse by scientific and educational bodies.¹⁰⁰ In short, unless the bill expressly and adequately immunizes traditional scientific and technical pursuits, the only limit on the database providers in most instances is what a monopoly market will bear.

C. Long-term Implications of the *Sui Generis* Model

We believe that the long-term implications of the proposed regime are potentially very damaging for science and technology. All science operates on databases. The near-complete digitization of data collection, manipulation, and dissemination over the past thirty years has ushered in what many regard as the transparency revolution.¹⁰¹ Every aspect of the

Phelps, Provost, University of Rochester, for AAU, ACE, and NASULGC) (“[E]xception for non-profit educational activities contains a broad, vague condition that vitiates its protection.”) [hereinafter Testimony of Phelps].

97. *See id.*

98. *See, e.g., Hearings, supra* note 57 (testimony of Joshua Lederberg, Nobel laureate, on behalf of NAS, NAE, IOM and American Association for the Advancement of Science (AAAS)) [hereinafter Testimony of Lederberg].

99. *See infra* text accompanying notes 151-55.

100. *See, e.g., Reichman & Franklin, supra* note 9, at 947-51; McManis, *supra* note 4.

101. *See, e.g., Paul F. Uhlir, From Spacecraft to Statecraft: The Role of Earth Observation Satellites in the Development and Verification of International Environmental Protection Agreements*, 2 GIS LAW 1 (1995). While we emphasize the impact of database protection on science and technology in this article, we predict that the larger economy will likewise suffer if the anticompetitive effects we foresee should materialize. There is hardly any sector of the economy that is not significantly engaged in the creation and exploitation of digital databases, and many—such as insurance, banking, or direct marketing—are completely database-driven. *See, e.g., Hearings, supra* note 57 (statement of

natural world, from the nano-scale to the macro-scale, all human activities, and indeed every life form, can now be observed and captured as an electronic database.

According to Nobel laureate Joshua Lederberg,

[d]ata are the building blocks of knowledge and the seeds of discovery. They challenge us to develop new concepts, theories, and models to make sense of the patterns we see in them. They provide the quantitative basis for testing and confirming theories and for translating new discoveries into useful applications for the benefit of society. They also are the foundation of sensible public policy in our democracy. The assembled record of scientific data and resulting information is both a history of events in the natural world and a record of human accomplishment.¹⁰²

1. Reversing the Transparency Movement

Science builds on science. In all areas of research, the collection of data sets is not an end in itself, but rather a means to an end, the first step in the creation of new information, knowledge, and understanding. As part of that process, the original databases are continually refined and recombined to create new databases and new insights. Typically, each level of processing adds value to an original (raw) data set by summarizing the original product, synthesizing a new product, or providing an interpretation of the original data.¹⁰³

The processing of data leads to a not readily apparent paradox. The original unprocessed, or minimally processed, data are usually the most difficult to understand or to use by anyone other than the expert primary user. With every successive level of processing, the data tend to become more understandable and frequently are better documented for the nonexpert user. As the data become more highly processed, documented, and formatted for easier use, they also are more likely to attract copyright protection.¹⁰⁴

Yet, it is the raw, noncopyrightable data that are typically of greatest use and value to researchers, who can manipulate and experiment with the

the Computer & Comm. Industry Assoc. of America and the Online Banking Association) (stressing how H.R. 354 threatens "legitimate reuse of information") [hereinafter Statement of CCIA et al.].

102. Testimony of Lederberg, *supra* note 98, at 5.

103. NATIONAL RESEARCH COUNCIL, PRESERVING SCIENTIFIC DATA ON OUR PHYSICAL UNIVERSE 16 (1995).

104. *Id.*; see also *supra* note 25.

original measurements in pursuit of their own research goals. If strong intellectual property protection of noncopyrightable data sets, which previously had the least commercial marketability, weakened the still nascent impetus toward transparency, it could disproportionately affect the availability of data most commonly used in basic research and higher education.

2. *Transaction Costs Unlimited*

The success of the U.S. basic research and educational system is predicated on the relatively unfettered access to and use of factual information; on a robust public domain for data; and on easy re-use, compilation, and value adding applications of data. Practically all databases developed in the pursuit of basic research and education are motivated by non-economic incentives such as the desire to create knowledge, the thrill of discovery, and the enhancement of professional status.¹⁰⁵ The new database laws, however, place an overriding emphasis on protecting original investments and on augmenting purportedly necessary economic incentives to create new databases. At the same time, they undervalue the adverse effects on scientific and technical progress, as well as the aggregate economic and social costs inherent in restricting and discouraging the downstream applications and transformative uses of noncopyrightable databases in general.¹⁰⁶

The lack of any restraints on licensing, especially on sole-source data providers, adds to the dangers inherent in the creation of a strong exclusive property right in collections of data.¹⁰⁷ In particular, the ability of data providers to override by contract even the limited exceptions that the new law may grant to public-interest users, including scientists, engineers, and educators, is of great concern. Without a concomitant duty to deal fairly and reasonably with public-interest users, these combined powers could lead to high prices for data and to the imposition of harsh and oppressive

105. *Cf., e.g.*, BITS OF POWER, *supra* note 69, at 113 (comparing contributors and users of scientific data to non-market models of "a family or clan, in which exchange is not monetized but depends on social norms specifying expected and well-understood levels of contribution").

106. *See, e.g.*, Testimony of Phelps, *supra* note 96, at 3-4 ("The Academic Environment and Activities threatened by H.R. 354").

107. *See, e.g.*, Ginsburg, *supra* note 7, at 175 ("When the data ... [are] not available elsewhere ... the potential breadth of the potential market is very troublesome.").

terms concerning both access and subsequent uses of data that would especially disadvantage academic researchers.¹⁰⁸

Moreover, scientists and engineers will have to defray increased transactional and administrative costs engendered by the need to enforce the different legal restrictions on newly obtained data, to institute new administrative guidelines regulating institutional acquisitions and uses of such data, and by associated legal fees. Because universities and government agencies are inherently conservative, risk-averse institutions, they will err on the side of caution and place additional limits on what researchers and educators can do in acquiring and using data in order to avoid the possibility of costly litigation.¹⁰⁹

The proposed database law would severely discourage the re-use, re-compilation, and other value adding uses of data. Anytime someone uses data in a "collection of information" protected by the proposed law, that user becomes exposed to claims that he or she will have harmed the database originator's actual or potential markets.¹¹⁰ As a practical matter, this means that once public-domain data are collected and used for one purpose, such as to prepare a compilation of poisons and antidotes, it will foster a strong disincentive to use the same data for other purposes lest those uses violate the "harm to other markets" principle. By the same token, database recompilers or value adders incur the risk of lawsuits for infringement every time their new database resembles some pre-existing database, whether those data were used or not.¹¹¹

One of the most serious problems of all is the risk of inhibiting the creation and exploitation of multiple-source data products, which have be-

108. See, e.g., Testimony of Lederberg, *supra* note 98, at 2-6 ("Progress in the creation and reuse of new knowledge for the national good depends on the full and open availability of government and government-funded data, and on fair and equitable availability of data from the private sector."); Reichman & Franklin, *supra* note 9, at 943-47.

109. For the problems that university administrators already foresee, see Testimony of Phelps, *supra* note 96.

110. See H.R. 354, 106th Cong. § 1402 (1999). However, unauthorized extractions or uses for *nonprofit* educational, scientific, or research purposes incur liability only for harm to actual markets. See *id.* § 1403(a)(1).

111. See, e.g., Statement of CCIA et al., *supra* note 101; Ginsburg, *supra* note 85, at 23-24 (stressing need for publisher to identify value-added contributions). The exception that permits anyone to make use of "insubstantial parts" of a collection of information is vitiated by the language inflicting liability for harm to the investor's "actual or potential market." See *supra* notes 53-55 and accompanying text. Because the user cannot know such matters in advance, the "potential harm" test emasculates the "insubstantial parts" exception in practice.

come the scientific method of choice for addressing hard new problems. Because research is increasingly conducted by teams, often operating from different institutions, the pertinent data “are drawn from multiple sources, recombined and merged with new data to produce data sets that may lead to new and unanticipated findings.”¹¹² As Joshua Lederberg testified at a hearing on H.R. 354, the “recent advent of digital technologies for collecting, processing, storing, and transmitting data has led to an exponential increase in the size and number of databases created and used. A hallmark trait of modern research is to obtain and use dozens or even hundreds of databases, extracting and merging portions of each to create new databases and new sources for knowledge and innovation.”¹¹³

In this regard, the Administration itself predicted that, under the current proposals, scientists and engineers would face rising transaction costs when attempting to create complex databases from multiple public and private sources. Also predicted are higher costs due to the burdens of administering national data centers and of carrying out related, large-scale management activities that currently benefit from the policy of open and unrestricted access to scientific and technical data.¹¹⁴

3. *Endless Monopolies and Diminished Access to Government Data*

Because many data providers are sole-source and an exclusive property right would greatly strengthen the legal and economic protection of these mini-monopolies, the proposed legislation seems likely to raise the costs of data acquisitions to researchers and educators generally, not to mention other consumers. Those costs would either be passed on to the government and the taxpayer through increased research contract and grant requests, or they would simply diminish the resources available to researchers and education. If the costs and restrictions on all downstream or transformative data users—whether in the public or private sector—similarly increased (as feared by database proprietors opposed to strong

112. Testimony of Phelps, *supra* note 96, at 3.

113. Testimony of Lederberg, *supra* note 98, at 6. *See also* Testimony of Phelps, *supra* note 96, at 3 (“In the academic community, ... databases are dynamic instruments: they are not only sources of information, but they themselves—or components of them—become ingredients in new products, both through the combination of multiple contemporaneous data sets to produce qualitatively new products, and through the re-analysis of prior data from new perspectives provided by new findings or new analytic tools.”).

114. *See* Letter from Andrew J. Pincus, General Counsel of the U.S. Dept. of Commerce, to Senator Patrick J. Leahy, Ranking Minority Member of the Senate Committee on the Judiciary (Aug. 4, 1998) (on file with authors) [hereinafter Administration’s Letter 1998].

ilarly increased (as feared by database proprietors opposed to strong protection), it would discourage socially and economically beneficial forms of exploiting factual data that have up to now been available from the public domain.¹¹⁵

The fifteen year term (which is potentially much longer because of loopholes favoring constantly updated, dynamic databases) is particularly likely to hamper the progress of science and technology, a prospect that troubles the Administration, too.¹¹⁶ Such long delays in unfettered access to and use of data will undermine the value of many data sets for most fields of research, including research pertaining to the formulation of government policy; and in other cases it will effectively remove them from comparative analysis with other, openly available, concurrent data sets.¹¹⁷

A fifteen year period appears completely arbitrary and has not been seriously compared with other, potentially shorter, periods of protection.¹¹⁸ The proposed legislation thus defeats a primary, constitutionally mandated purpose of intellectual property laws, which is to establish a public domain that "promotes science and the useful arts,"¹¹⁹ from which researchers, educators, and other downstream users can build on previous contributions to further knowledge.

The proponents of the legislation say that nothing prevents a user or competitor from independently creating an equivalent database.¹²⁰ But many databases cannot be recreated from scratch. Data that are time-sensitive, unique, very old, or prohibitively expensive fit this description. In research, this includes virtually all observational data sets of transient natural phenomena, as well as data from very costly or labor-intensive experiments. Furthermore, a basic underlying principle in research and edu-

115. See Statement of CCIA et al., *supra* note 101, at 8-10 (noting that some commercial database providers oppose strong protection owing in part to fears about consequences of sole-source market structure).

116. See *supra* note 55 and accompanying text; Statement of Pincus, *supra* note 87, at 21-27.

117. Besides retarding scientific and technical progress, the fifteen-year term has no apparent justification in the rapidly moving commercial database industry either, where economic exploitation of most data products is typically measured in months and years, and even minutes and hours, rather than decades. See Statement of CCIA et al., *supra* note 101, at 10.

118. See, e.g., Ginsburg, *supra* note 85, at 30 (acknowledging soundness of lead-time criterion not adopted in E.U. Directive); Reichman & Samuelson, *supra* note 7, at 145 (proposing lead-time criterion and relating it to lack of actual or legal secrecy).

119. U.S. CONST. art. I, § 8, cl. 8.

120. See, e.g., Ginsburg, *supra* note 7, at 175.

ation is that the creation of new knowledge should build on the base of existing data and information, and that scientists and engineers should not have to duplicate previous factual compilations or discoveries in socially and economically inefficient ways.¹²¹ Protection of investments in factual databases is not the only interest that the law should seek to protect in this area.

Public interest users in the United States are likewise concerned about the applicability of the E.U. Database Directive to government data and the potential restrictions on access to and use of European public-sector data. Moreover, even though the proposed U.S. legislation does expressly exempt government data from its scope of protection, there are concerns that, as drafted, this exemption could be circumvented in several ways.¹²²

This can occur, for example, if the contractors or grantees are not expressly required either to provide their data back to the government for public dissemination, or to make the data publicly available themselves under appropriate terms and conditions.¹²³ Absent such universal vigilance by the government, a lot of data produced as a direct result of public funding could end up under proprietary control of researchers or their institutions. Because most of the noncopyrightable databases generated with government funding in the United States are actually created by non-government employees, whether in academia or industry, the failure of government agencies to enforce this exemption could have a far-reaching impact on the full and open availability of publicly funded data. Indeed, there is some risk that government agencies could increasingly view database protection as an income-generating opportunity, like their European

121. See, e.g., Statement of Lederberg, *supra* note 98, at 4-5; Testimony of Phelps, *supra* note 96, at 3-4; Statement of Neal, *supra* note 96.

122. See H.R. 354, 106th Cong. § 1404 (1999); see also, Statement of Pincus, *supra* note 87, at 13-20 ("Third Principle—Preserve access to government data"). The federal basic research budget alone is estimated to be over \$19.5B/year, of which a sizeable fraction is devoted to the creation, maintenance, dissemination, and analysis of scientific and technical data. See INTERSOCIETY WORKING GROUP, AMERICAN ASS'N FOR THE ADVANCEMENT OF SCIENCE REPORT XXIV: RESEARCH AND DEVELOPMENT FY 2000 71 (1999). Moreover, the government at all levels produces data of other importance to the nation, including economic growth, public health and safety, regulatory requirements, cultural affairs, and many other functions. All citizens have an interest in preserving full and open access to all government data that are not otherwise restricted by national security, privacy, or other legitimate limitations.

123. See, e.g., Ginsburg, *supra* note 85, at 28 ("In the case of sole-source government information, public domain policy favors making the information available to market rivals.").

counterparts.¹²⁴ As more university research is funded by private sources, more data will likely be removed from the public domain in the form of income-producing products.

Still other legislation, if combined with increased database protection, could further limit the principle of full and open access to government data. For example, the Commercial Space Act of 1998 encourages NASA to purchase space and earth science data collection and dissemination services from the private sector and to treat data as commercial commodities under federal procurement regulations.¹²⁵ When coupled with strong protectionistic measures, such as those contemplated by H.R. 354, we could eventually witness the passing of substantial amounts of data from the public domain of entire federal agencies. It also remains unclear if the government concludes an arrangement with a private sector party to disseminate public data or information, whether there will be adequate safeguards that either promote competition or that require low-cost access for public-interest users.¹²⁶

4. *Gaming the Cooperative Ethos*

Finally, a high-protectionist regime tends to undermine scientific and technical cooperation over time and to exert a progressive chilling effect on data-intensive research. As scientists, engineers, and their employing institutions become more accustomed to a new legal regime that encourages the commercial exploitation of their own research data sets, the cooperative culture that has become the hallmark of so many fields of science will be threatened.¹²⁷ Universities have already indicated that they intend to commercially exploit databases, and they have obtained an exemption for state universities from the government data exception in the proposed legislation.¹²⁸ If scientific institutions in one segment of the research community try to commercially exploit their colleagues in other institutions or countries, still others will be tempted either to emulate such be-

124. See *supra* note 88 and accompanying text.

125. See Commercial Space Act of 1998, Pub. L. No. 105-303 (1998).

126. Cf. Statement of Pincus, *supra* note 87; Ginsburg, *supra* note 85, at 27-29.

127. See, e.g., BITS OF POWER, *supra* note 69, at 169-70 (urging the scientific community to organize its own administration of data in order to preserve the sharing ethos).

128. See H.R. 354, 106th Cong. § 1404(a)(1) (1999) (limiting the term of protection); Testimony of Phelps, *supra* note 96, at 17 (noting that "universities and colleges are not only users of compilations of information, they also act as creators of collections that should be [entitled to] protection to the same extent as collections created by commercial providers").

havior or to cut off cooperation. Either way, science and technology would suffer.

Even if scientific data exchanges in established cooperative research programs were allowed to continue among a select group of principal investigators and an approved class of associated researchers, it would become increasingly difficult for other researchers outside the officially sanctioned group to obtain full and open access to the program data. This result, of course, would discourage interdisciplinary research and applications, contrary to the interests of technological innovation and the advancement of knowledge.

If simple exchanges of data and access to single databases became legally threatening or prohibitively expensive, imagine the potential transactional burdens that ill-conceived laws could impose on data compilers or users who needed to integrate data from multiple, or even hundreds, of different sources. This brings us to what may well be the most profound—and insidious—impact of the proposed legal regime on science and technology: the lost opportunity costs that will be repeated thousands of times each day across the basic and applied research communities. If scientists and engineers must choose between spending a lot of administrative time and a larger percentage of their valuable research grants on acquiring data and doing other, less data-intensive work, they will increasingly opt for the second route, despite the astounding yields that have so far been harvested from data-intensive research under existing conditions.

For all these reasons, an overly protective database regime would seriously impede the use, reuse, and transformation of the factual data that are the lifeblood of science and technology. In a worst case scenario, this law would first disrupt the system of cheap access to upstream data for purposes of basic research, it would then lead to ever higher prices for the acquisition of data used in applied research, and finally, it would strangle the ability of value-adding researchers and industries to improve, transform, or develop follow-on databases and related information products.¹²⁹ These outcomes would, in turn, greatly reduce the downstream applications of scientific breakthroughs subject to exclusive property rights. In sum, putting a strong property right too far upstream too soon¹³⁰ could have a disastrous effect on the long-term competitiveness of the U.S. economy and

129. See, e.g., Statement of CCIA et al., *supra* note 101.

130. Cf. Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCIENCE 698-701 (1998).

would undermine a key comparative advantage this country enjoys in the high-tech sectors of the global marketplace.

III. RECENT DEVELOPMENTS: THE QUEST FOR AN APPROPRIATE UNFAIR COMPETITION APPROACH

Concern about these issues mobilized the scientific, educational and library communities to express their views both to Congress and to the World Intellectual Property Organization ("WIPO"). For example, the Presidents of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine sent several letters to the Administration and to leading members of Congress responsible for this legislation.¹³¹ The International Council for Science ("ICSU") likewise intervened at relevant meetings of WIPO, and documents submitted by ICSU have played a prominent role in successful efforts to block rapid or premature efforts to launch an international treaty regulating databases modeled on the unbalanced E.U. Directive.¹³² ICSU has also begun direct consultation with publishers' representatives, with a view to working out some common understanding applicable to database protection issues affecting the scientific community.

Initial efforts in 1997 and early 1998 to slow the legislative process in the U.S. House of Representatives were not successful. The concerns of the scientific community were largely ignored by the House Committee on the Judiciary's Subcommittee on Courts and Intellectual Property, which first pushed H.R. 2652 (the "Collections of Information Antipiracy Act") through the House in the spring of 1998 and then had that bill attached to the Digital Millennium Copyright Act, H.R. 2281, in July 1998.¹³³

131. Letters from the three Academy Presidents to Mickey Kantor, Secretary of Commerce (Oct. 9, 1996) and to Senator Orrin Hatch, Chair of the Senate Committee on the Judiciary (July 10, 1998) (on file with authors).

132. See ICSU Position Paper, *supra* note 34; Reichman & Samuelson, *supra* note 7, at 95-102, 110-13 (describing efforts to have WIPO adopt a database protection treaty in 1996).

133. See H.R. 2652, 105th Cong. (1998); H.R. 2281, 105th Cong. (1998); see also COPYRIGHT OFFICE, REPORT ON LEGAL PROTECTION FOR DATABASES (1997), available at <<http://lcweb.loc.gov/copyright/reports>>. For the latest position of the Copyright Office, see *Hearings*, *supra* note 57 (statement of Marybeth Peters, Register of Copyrights) (noting improvements in H.R. 354, but stating that "it is important not to inhibit or raise the cost of public interest uses" and that an "appropriate statutory balance should result in optimizing the availability of reliable information to the public") [hereinafter Statement of Peters].

A. The Administration's Position

More recent developments, however, have been favorable to the interests of the scientific and educational communities. To begin with, an interagency review initiated within the Administration in the spring of 1998 produced a series of important position papers that supported the theses that the scientific community had already put forward. On August 4, 1998, the General Counsel of the U.S. Department of Commerce, Andrew Pincus, wrote Senator Patrick Leahy, Ranking Minority Member of the Senate Committee on the Judiciary, to advise that any new database legislation must avoid capture by private parties of government data and that "any effects [it may have] on non-commercial research should be *de minimis*."¹³⁴ This letter was sent as a consensus position of all departments and agencies of the Administration.

Consistent with these views, the Administration expressed concerns about possible

increase[d] transaction costs in data use, particularly where larger collections integrate data sets originating from different parties or where different parties have added value to a collection through separate contributions.... This is especially important for large-scale data management activities, where public investment has leveraged contributions from the private and non-profit sectors.¹³⁵

The letter went on to express further concerns "that the ... exception for noncommercial research and educational uses does not ensure that legitimate non-commercial research and educational activities are not disrupted by the prohibition against commercial misappropriation" and that sole-source providers might unduly burden "access and use" by this sector.¹³⁶ The Administration reiterated most of these same concerns in testimony before the House Subcommittee on Courts and Intellectual Property, at hearings concerning H.R. 354 on March 18, 1999.¹³⁷

The Administration's initial letter of August 4, 1998, also referenced the Department of Justice's "serious constitutional concerns that the First Amendment restricts Congress's ability to enact legislation" of this kind, and that other constitutional obstacles would have to be overcome.¹³⁸

134. Administration's Letter 1998, *supra* note 114.

135. *Id.*

136. *Id.*

137. See Statement of Pincus, *supra* note 87.

138. Administration's Letter 1998, *supra* note 114.

These constitutional impediments were elaborated in a 26-page memorandum by the Legal Counsel of the U.S. Department of Justice, dated July 28, 1998, which detailed a serious indictment of the *sui generis* model then pending before Congress.¹³⁹

On September 28, 1998, moreover, the Chair of the Federal Trade Commission ("FTC"), Robert Pitofsky, wrote the Chair of the House Committee on Commerce, Tom Bliley, to express additional concerns about the pending database legislation.¹⁴⁰ In particular, the FTC found that "certain provisions within the proposed legislation raise concerns about possible unintended, deleterious effects on competition and innovation," and that "the potential for anti-competitive use of a 'collection of information' is substantially increased when there is only a single source for the data."¹⁴¹

Finally, it should be noted that the United States Patent and Trademark Office ("USPTO") held a conference on April 28, 1998, to reexamine database protection and access issues. In July 1998, it issued a Report that, while endorsing a modified form of *sui generis* database protection, expressed support for some of the concerns that the scientific and educational communities had been voicing.¹⁴²

B. A Negotiated Discussion Draft in the Senate

In late July of 1998, the Chair of the Senate Committee on the Judiciary, Senator Orrin Hatch, invited representatives of some of the major stakeholder organizations and companies to participate in strenuous negotiations, which lasted from the beginning of August through early October. These negotiations were conducted under the leadership of Senator Hatch's Counsel for Intellectual Property, Edward Damich, and with the

139. See Memorandum from William Michael Treanor, U.S. Dept. of Justice, Office of the Deputy Ass't. Att'y. Gen., to William R. Marshall, Associate White House Counsel, "Constitutional Concerns Raised by the Collections of Information Antipiracy H.R. 2652" (July 28, 1998) (on file with authors) [hereinafter DOJ Memorandum]. See also Letter from Professor Marci Hamilton to Chairman Howard Coble (Feb. 10, 1998) (on file with authors) (detailing serious Constitutional concerns) [hereinafter Letter from Hamilton].

140. See Letter from Robert Pitofsky, Chair of the Federal Trade Commission, to Tom Bliley, Chair of the House Committee on Commerce (Sept. 28, 1998) (on file with authors).

141. *Id.* at 7, 14.

142. See U.S. PATENT AND TRADEMARK OFFICE, REPORT ON (AND RECOMMENDATIONS FROM) APRIL 1998 CONFERENCE ON DATABASE PROTECTION AND ACCESS ISSUES (1998).

participation of the counterpart staffer in Senator Leahy's office, Marla Grossman.

1. Clarifying the Demands on Scientific and Technical Users

The U.S. Academies took the unusual step of participating directly in these negotiations.¹⁴³ They submitted a series of alternative proposals aimed at providing a balanced piece of legislation that would protect publishers against free-riding conduct and preserve the incentive to invest (through a true unfair competition approach) without creating a strong exclusive property right in collections of data and factual information.¹⁴⁴

Although the direct negotiations produced no major breakthroughs or compromise solutions, they did succeed in clarifying the different positions. It seems fair to say that, when exposed to direct interrogation, the publishers' detailed demands more than justified the scientific and educational communities' initial concerns.¹⁴⁵ Indeed, in response to one hypothetical situation after another, the publishers' representatives made it clear that the exclusive property right they championed for the digital network system would in fact engender the kind of legal and contractual demands on scientific, technical, and commercial users of protected databases that critics of the proposed legislation had been fearing and that are described in this article.

143. The authors of this article represented the National Academy of Sciences (NAS), the National Academy of Engineering (NAE), and the Institute of Medicine (IOM) during these negotiations.

144. See, e.g., National Academy of Sciences et al., Proposed Amendments to H.R. 2281: Synopsis, Corrections, and Text, submitted for consideration by the Senate Committee on the Judiciary (Aug. 11, 1998) (unpublished manuscript, on file with authors) [hereinafter NAS, Synopsis]; National Academy of Sciences et al., Proposed Amendments to H.R. 2281: Explanatory Memorandum (Part I) (Aug. 13, 1998) (unpublished manuscript, on file with authors) [hereinafter NAS, Explanatory Memorandum]; NAS et al., Opponents' Revised Amendments to H.R. 2281 (Sept. 4, 1998) (unpublished manuscript, on file with authors) (concerning (1) permitted acts for scientific, educational, and research purposes; (2) exclusions; (3) definition of "collections of information"; and (4) licensing).

145. For the latest iteration of the publishers' position, see *Hearings, supra* note 57 (testimony of Daniel C. Duncan, Vice President for Gov't. Affairs, Software and Information Industry Assoc.) [hereinafter Testimony of Duncan]; *Hearings, supra* note 57 (testimony of Marilyn Winokur, Exec. Vice President, Micromedex for Coalition Against Database Piracy); *Hearings, supra* note 57 (testimony of the National Assoc. of Realtors). See also *Hearings, supra* note 57 (statement of Michael K. Kink, AIPLA); *Hearings, supra* note 57 (testimony of the Agricultural Publishers Assoc.).

Perhaps because the publishers' actual demands amply confirmed the concerns that the Administration's own position papers had expressed, the final phases of the negotiations, as mediated by the Senate staffers, produced far-reaching modifications to the database component of H.R. 2281. On January 19, 1999, Senator Orrin Hatch placed in the Congressional Record a statement acknowledging that "considerable progress" had been made during the aforementioned negotiations and that, "in the end we were close to a workable compromise."¹⁴⁶ Senator Hatch also put in the Record "a discussion draft that is identical to the last of the discussion drafts ... [he had] offered last year."¹⁴⁷

2. *Compromise Proposals*

The changes incorporated in the last Discussion Draft substantially reflected the Academies' own position. Although there is no way of knowing the degree of assent to all the various provisions it contained, it is worth reviewing the package of compromise proposals embodied in the last Hatch Database Discussion Draft of October 5, 1998.¹⁴⁸

First, the strong property right approach was nudged closer to a true "misappropriation" (unfair competition) approach. This was accomplished by conditioning liability on acts that "cause *substantial* harm to the actual or neighboring market" of database proprietors,¹⁴⁹ and by inviting courts, in the draft legislative history, to determine "substantial harm" in light of "whether the harm is such as to significantly diminish the incentive to invest in gathering, organizing, or maintaining the database."¹⁵⁰

Second, a full exception that would immunize customary scientific activities was adopted by the Senate staffers, in place of the limited and unacceptable "fair use" approach that the Administration had eventually rec-

146. See 106 CONG. REC. S.316 (daily ed. Jan. 19, 1999) (statement of Senator Hatch on Database Antipiracy Legislation) [hereinafter Statement of Hatch].

147. See *id.* at S.322-26 (Chapter 14—Protection of Databases) [hereinafter Hatch Database Discussion Draft].

148. See *id.* An abridged selection of these points was used in a letter from the Presidents of the NAS, NAE, and IOM to Howard Coble, Chair of the Subcommittee on Courts and Intellectual Property of the House Committee on the Judiciary, March 9, 1999, and subsequently in the testimony of Joshua Lederberg before the Subcommittee at its March 18, 1999 Hearing. See *supra* notes 57, 114

149. See Hatch Database Discussion Draft, *supra* note 147, § 1302.

150. Hatch Database Discussion Draft Proposed, Conference Report Language, § 1302, at 33 (on file with authors).

ommended.¹⁵¹ A “fair use” approach, modeled on copyright law, would fail because other basic copyright immunities and exceptions, especially the idea-expression dichotomy, would not carry over into the database protection environment. On the contrary, because a database law protects collections of facts and data that are ineligible under copyright laws (and because scientists perceive no valid distinction between “data” and a “collection of data” in a dynamic electronic database), basic research methods that were previously permissible would become infringing acts under such a law. The burden would then fall on scientists and engineers to show that a vague fair use exception should excuse some of these infringing acts from whatever test of harm was adopted.

In contrast, the Academies successfully argued that customary and traditional scientific activities should remain untouched and unhampered by any new database protection law, exactly as the government’s initial position paper had maintained.¹⁵² To this end, section 1304 of the final version of the Hatch Database Discussion Draft stated the following proposition:

Nothing in this chapter shall prohibit or otherwise restrict the extraction or use of a database protected under this chapter for the following purposes:

- 1) for illustration, explanation, or example, comment or criticism, internal verification, or scientific or statistical analysis of the portion used or extracted; and
- 2) in the case of nonprofit scientific, educational, or research activities by nonprofit organizations, for similar *customary or traditional purposes*.¹⁵³

Only if scientists, engineers or educators working at nonprofit organizations caused substantial harm to the database-maker by using unreasonable and non-customary amounts of the collection for a given purpose, or if they in fact produced a market substitute for the original, or otherwise sought to avoid paying for the use of research tools devised as such, would

151. See Administration’s Letter 1998, *supra* note 114; see also Memorandum from the Administration to the Senate Committee on the Judiciary, *Senate Draft Developments and Suggested Additional changes to Address Key Concerns for Discussion* (rev. Sept. 30, 1998) (on file with authors).

152. See Administration’s Letter 1998, *supra* note 114. However, the Government’s latest position paper leans towards the compromise “fair use” provision introduced in H.R. 354, which we judge to be inadequate. See Statement of Pincus, *supra* note 87, at 29-31.

153. Hatch Database Discussion Draft, *supra* note 147, § 1304(a) (emphasis added).

liability kick in.¹⁵⁴ On this approach, the burden fell on publishers to show that scientists had crossed the line of permitted, traditional, or customary uses, which were otherwise immunized. The guiding principle that science, technology, and education should be left no worse off after enactment than they were before, as proposed by ICSU,¹⁵⁵ would thus have been implemented.

Third, additional immunities and exceptions favoring certain instructional and library uses of databases were also defined,¹⁵⁶ although more thought needs to be given to educational users generally in this context.¹⁵⁷ Fourth, efforts were also made to reduce the likelihood that private interests might permanently capture government-generated data,¹⁵⁸ although more remains to be done on this score as well.¹⁵⁹

Fifth, a clearly-worded duration clause ending protection after fifteen years reduced (but did not altogether eliminate) the risk of perpetual protection.¹⁶⁰ A rudimentary database deposit scheme was also proposed, which increased the likelihood of data eventually entering the public do-

154. *See id.* § 1304(b):

In no case may a use or extraction for a purpose described in subsection (a) be permitted if the substantial harm referred to in Section 1302—

1. arises because the amount of the portion used or extracted is more than is reasonable and customary for the purpose;

2. consists of the use or extraction being intended to, or being likely to serve as a substitute for or to supplant all or a substantial part of the database from which the extraction or use is made or an adaptation thereof that is protected under this chapter;

3. arises because the extraction or use is intended to avoid payment of reasonable fees for use of a database incorporated into a product or service specifically marketed for educational, scientific, or research purposes; or

4. arises because the use or extraction is part of a pattern, system, or repeated practice by the same party, related parties, or parties acting in concert with respect to the same database or a series of related databases.

155. *See* ICSU Position Paper, *supra* note 34, at 2.

156. *See* Hatch Database Discussion Draft, *supra* note 147, § 1307.

157. *See also* Statement of Phelps, *supra* note 96, at 16 (proposing exemption for non-profit teaching activities).

158. *See* Hatch Database Discussion Draft, *supra* note 147, § 1305(a)-(c).

159. *See* Statement of Pincus, *supra* note 87, at 13-20 (detailed proposals concerning government-funded data and the need to avoid its "capture" by commercial interests); Ginsburg, *supra* note 85, at 27-29.

160. *See* Statement of Pincus, *supra* note 87, at 24-27.

main, albeit in a cumbersome and, perhaps, costly fashion.¹⁶¹ If that route were taken, more incentives would be needed to ensure that deposits were actually made. However, the Administration favors developing other, simpler incentives to ensure the availability of public domain data that are worth exploring.¹⁶²

Sixth, the need for some regulation of licensing terms and conditions was expressly recognized. A series of provisions required periodic studies of the misuse doctrine as applied to licensing agreements, or to the use of technological measures that might frustrate the "permitted acts" clause of the bill. Particular grounds for the study included sole-source provider contracts that imposed unreasonable terms or conditions; tying or other practices traditionally recognized as abusive; and practices shown to have "prevented access to valuable information for research, competition, or innovation purposes."¹⁶³ The draft legislative history then clarified that courts were free to apply these same criteria to claims of misuse arising after the time of enactment and need not "refrain from applying the doctrine of misuse until the study is completed."¹⁶⁴ There was some further possibility that criteria for evaluating the misuse of licensing agreements might have ultimately been codified in the operative clauses of the Act itself.

Finally, the draft legislative history to accompany these measures also clarified the definition of databases in ways that tended to exclude ordinary literary works, and it denied protection "to any ideas, facts, procedure, process, system, method of operation, concept, principle, or discovery, as distinct from the collection that is the product of investment protected by this Act."¹⁶⁵ Needless to say, we think the proposed database law should expressly codify these provisions. Indeed, the fact that the bill's proponents oppose inclusion of such a basic limitation in H.R. 2281 only serves to reinforce our concerns about the true nature and extent of their intended exploitation of the legislation's most restrictive provisions.

161. See Hatch Database Discussion Draft, *supra* note 147, § 1310(d); Statement of Pincus, *supra* note 87, at 24-27.

162. See Statement of Pincus, *supra* note 87, at 24-27 (discussing dated identity tags and statutory defenses to liability, including lack of public availability of government-funded data); see also Ginsburg, *supra* note 85, at 24.

163. Hatch Database Discussion Draft, *supra* note 147, § 4. See also *id.* § 1306(a) (preserving doctrine of misuse). But see *id.* § 1306(e) (allowing licensors unrestricted freedom of contract).

164. Hatch Database Discussion Draft, Proposed Conference Report Language, *supra* note 150, at 36-37.

165. *Id.* at 131.

C. Uncertain Future of the Database Protection Law

The foregoing discussion reveals the extent to which the Hatch Database Discussion Draft evolved away from the strong exclusive property right approach, adopted in the E.U. Directive, toward a more balanced unfair competition approach that protected publishers against piracy while consciously avoiding harm to science, education, and other public-good uses of data. Of course, not all the issues of concern to science were addressed in a fully satisfactory manner; but given the need for compromise and consensus, the ability of the staff to produce a relatively balanced bill from such unpromising material as the House bill deserves commendation.

Perhaps the biggest unaddressed issue was that of value-adding uses. The Discussion Draft did not resolve the tensions between a dominant group of database publishers, who seek to control value-adding uses of protected collections, and a dissident group of publishers and allies, who believe value-adding uses should remain as unfettered as possible.¹⁶⁶ On this point, the Academies proposed a scheme favoring easy use of data for commercial value-adding purposes in exchange for the payment of reasonable royalties under an automatic licensing scheme,¹⁶⁷ but neither side would accept this approach. Nevertheless, under the misappropriation approach to "substantial harm," as elaborated in the Draft Legislative History,¹⁶⁸ courts could work out the criteria for balancing incentives to invest against incentives to compete for the short run, and these case-by-case solutions could be legislatively evaluated later on.

In the end, the Hatch Discussion Draft was not adopted mainly because time ran out in which to remove the last remaining wrinkles that prevented an agreed compromise.¹⁶⁹ As a result, the database component of H.R. 2281, Title V, was stripped from the Digital Millennium Copyright Act, which was enacted at the end of the legislative year. Work on

166. Compare, e.g., Testimony of Duncan, *supra* note 145, at 9-10 (presenting a case for broad scope of protection) with Testimony of CCIA et al., *supra* note 101, at 6-8 (advocating less restricted commercial reuse of data).

167. See NAS, Synopsis, *supra* note 144, § 1303(g) (distinguishing use by competitors on distant markets from use in direct competition for these purposes). See also Reichman & Samuelson, *supra* note 7, at 146 (proposing "automatic license built into ... [a] modified liability right itself"); Ginsburg, *supra* note 85, at 28 (proposing automatic license for sole-source collections of government data to benefit "market rivals").

168. See *supra* note 150.

169. See *supra* note 146 and accompanying text (expressing Senator Hatch's belief that considerable progress had been made).

database protection has begun all over again under the aegis of a new Congress, which convened in January 1999.¹⁷⁰

During the fall of 1998, some members of the coalition that had opposed H.R. 2281 drafted still another bill that sought to implement unfair competition principles more aggressively than was contemplated in the final Hatch Discussion Draft.¹⁷¹ There exists some support for this so-called "minimalist" unfair competition approach, which would protect databases only against wholesale duplication for an indefinite period of time.¹⁷² However, this solution could easily degenerate into a *de facto* exclusive property right conferring perpetual protection by the back door.

Whatever happens next, the final version of the Hatch Discussion Draft constituted a milestone along the route towards a more balanced model of database protection, and its lessons should inform the next round of legislative deliberations. There is unofficial and anecdotal evidence that the Japanese government may also embark upon a true unfair competition approach, which, if true, would afford a unique opportunity for the United States and Japan to present a united front to the rest of the world. In that event, other countries would probably move in the direction of a more balanced unfair competition regime, which might leave the E.U. alone to continue its experiment with a strong property right, or to modify its Directive so as to obtain a more balanced system of protection with fewer social costs.

Unfortunately, the scientific community will experience serious challenges to the policy of easy access to, and unrestricted uses of data, regardless of the approach to database protection that ultimately emerges from Congress and the legislation of other countries. As pointed out at the beginning of this article, publishers can already control the dissemination of data by combining technical protection measures with adhesion contracts in the online environment, even without the adoption of specific database legislation.¹⁷³ While the presence of an intellectual property right would strengthen the publishers' position and put the scientific and technical communities under grave legal disadvantages,¹⁷⁴ the absence of an ex-

170. See H.R. 354, 106th Cong. (1999); *supra* text accompanying notes 57-61.

171. See Statement of Hatch, *supra* note 146, at S.320-22 (Proposed Bill to Amend Title 17, United States Code, to Promote Research and Fair Competition in the Database Industry) [hereinafter Minimalist Protection Bill of 1999].

172. See Minimalist Protection Bill of 1999, *supra* note 171, §§ 1401, 1408.

173. See *supra* text accompanying notes 10-18.

174. See Reichman & Franklin, *supra* note 9, at 911-14, 947-51 (exploring impact of online adhesion contracts with or without a codified property right underneath).

clusive property right would not free them from the need to rethink their whole approach to maintaining the unrestricted flow of scientific and technical data in an emerging information economy.

We trust that two new studies by the National Research Council, which are examining these issues, will shed further light on the options for science when they are published in mid-1999. Meanwhile, it seems clear that the scientific and technical communities will have to consider ways of reconciling a greater degree of commercialization for databases generated within the academic community with the need to maintain privileged access to the same databases for scientific and other public interest objectives. Universities and research institutions that generate data will thus have to develop rules for disciplining grants and the uses of data obtained from grants. Separate channels for the nonprofit distribution of scientific and technical data may have to be created, with particular rules for participating organizations. Ultimately, it may also prove desirable to develop an extended licensing authority for certain classes of scientific data, in order to administer these resources with low transaction costs and uniform rules for commercial and non-commercial users.¹⁷⁵

In general, efforts must be made to preserve the sharing ethos with respect to publicly-generated scientific data, to encourage those who invest in the production of privately-generated data to provide price discrimination in favor of the scientific and educational communities, and to develop differentiated products for the non-profit sector.¹⁷⁶ As the Academies recently explained, Congress should strive to reconcile legitimate measures to repress parasitical copying of protected databases with the equally legitimate needs of the scientific, technical, and educational communities. These communities require:

- access to data on fair and reasonable conditions;
- the ability to use the data accessed for research or educational purposes; and
- freedom from contractual or technical interference with these pursuits.¹⁷⁷

These objectives will, in turn, require close collaboration with governments. The goal is to ensure that data generated at the taxpayers' ex-

175. Cf. Merges, *supra* note 1.

176. See BITS OF POWER, *supra* note 69, at 166-71.

177. See NAS, Explanatory Memorandum, *supra* note 144, at 9.

pense remain available at least for scientific and educational purposes, and that efforts to stimulate greater investment in the development of new databases do not end by creating barriers to entry or otherwise discouraging follow-on innovation and public good uses of the building blocks of knowledge.

IV. PRESERVING THE CONSTITUTIONAL BALANCE OF INTERESTS IN THE NETWORKED ENVIRONMENT

Because everything on the Internet is potentially a “database” or a “collection of information” in our increasingly information-based economy, the law that protects collected information will determine the level of competition and prices in that economy. The EU Directive—and to almost the same extent its counterpart proposal pending in the United States¹⁷⁸—opt for a very high level of protection. These regimes buttress mini-monopolies of data and information that could threaten the advance of scientific and technical research, hinder the creation of legitimate new commercial information products, and hurt downstream consumer interests.

A. The Competitive Ethos Under Attack

The fallacy behind most proposals for strong forms of database protection is that they ignore the dual nature of data and information as such. On one level, data function as a raw material of the information economy, a basic ingredient of the public domain, from which scientists and entrepreneurs both draw to fashion their respective products. On a second level, data and information are bundled into downstream products that attract intellectual property rights and related contractual agreements. The mistake is to presume that strong intellectual property rights that were empirically well-suited to downstream applications—mainly derived from the patent and copyright models—are equally well-suited to upstream regulation of the data as inputs into the process of innovation.

The opposite is true. If we balkanize the public domain and make the transaction costs of recreating it by contracts prohibitively expensive and complex, a dysfunctional legal system will impede the cumulative and sequential development of technical paradigms by depriving routine innovators of access to the building blocks of knowledge.¹⁷⁹

178. See E.U. Directive, *supra* note 6; H.R. 354, 106th Cong. (1999).

179. See Heller & Eisenberg, *supra* note 130; Suzanne Scotchmer, *Standing on the Shoulders of Giants: Cumulative Research and the Patent Law*, 5 J. ECON. PERSPECTIVES 29 (1991); Reichman & Franklin, *supra* note 9, at 884-88. Cf. generally Michael A. Hel-

The truth is that traditionally we have left small grain-size innovation to weaker forms of entitlement, that is, to liability principles rooted in unfair competition law, rather than strong property rights, and this has been a basic premise on which the competitive economy of the industrial revolution was constructed.¹⁸⁰ These lessons are still germane to the information economy—lessons sounding in reverse engineering and the reuse of ideas, rather than in legally supported monopolies on products of routine innovation and investment.

Until convincing evidence to the contrary accrues, we should address the risk of market failure in the information economy by erring on the side of underprotection rather than overprotection. This follows because there is no real or potential shortage of investment in this milieu once the causes of market failure are controlled; and it is sound public policy, because we do not wish needlessly to encourage the monopolization of the sources of factual data, to deter value-adding innovators, or to retard the progress of science.¹⁸¹

B. The Constitutional Dilemma

The inclination to place strong intellectual property rights in upstream collections of information is contrary to our entire intellectual property tradition and to our basic constitutional heritage. For some forty years, the late Professor Melville Nimmer, a leading authority on both copyright and First Amendment law, taught that copyright protection would violate First Amendment guarantees of free speech were it not for the judicial exclusion of ideas and facts from the reach of the exclusive property rights granted to authors and artists.¹⁸² In 1976, Congress codified that exclusion

ler, *The Tragedy of the Anticommons: Property in the Transition from Marx to Markets*, 111 HARV. L. REV. 621 (1998) (showing that too many rights to exclude produce anti-commons effects as bad as the lack of any powers to exclude).

180. See generally Reichman, *Legal Hybrids*, *supra* note 3, at 2434-44; Reichman, *Charting*, *supra* note 3, at 485-96, 504-17.

181. See generally J.H. Reichman, *Solving the Green Tulip Problem: Repackaging Rights in Subpatentable Innovation*, paper presented to New York University Conference on "Intellectual Products: Novel Claims to Protection and Their Boundaries," Engelberg Center on Innovation Law and Policy, La Pietra, Italy (June 25-28, 1998) (unpublished manuscript, on file with authors).

182. See M.B. NIMMER & D. NIMMER, *NIMMER ON COPYRIGHT* § 1.10 (1998); R.A. SMOLLA, *SMOLLA & NIMMER ON FREEDOM OF SPEECH* § 21:8 (1998); Robert C. Denicola, *Copyright and Free Speech: Constitutional Limitations on the Protection of Expression*, 67 CALIF. L. REV. 283 (1983); M.B. Nimmer, *Does Copyright Abridge the First Amendment Guarantees of Free Speech and Press?*, 17 UCLA L. REV. 1180 (1970); Al-

in Section 102(b) of the General Revision of Copyright Law,¹⁸³ and in 1991, the Supreme Court, in *Feist Publications, Inc. v. Rural Telephone Service, Co.*,¹⁸⁴ reconfirmed the constitutional prohibition against an exclusive property right in either facts or ideas.

Proponents of H.R. 354 and its predecessors openly concede that the "harm to actual or potential markets" test was drawn from Section 107(4) of the 1976 Copyright Law, which codified its fair use provisions.¹⁸⁵ This is a constitutionally dubious admission because the very purpose of Section 107(4) is to confirm that protection of the author's market interests in both primary and secondary markets constitutes the true goal of the copyright law's exclusive rights, exactly as Judge Frank declared in his famous opinion in *Arnstein v. Porter*.¹⁸⁶

When transplanted to the database milieu, however, the protection of mere investment in databases that do not rise to the level of creative works of authorship against "harm to actual or potential markets"¹⁸⁷ indirectly creates an exclusive property right in noncopyrightable collections of data, which governs both primary and secondary markets. Once collected, no one can make further use of the facts and data contained in the collection without the compiler's permission, even though Section 102(b) of the Copyright Law states that facts and ideas are not fit subjects of an exclusive property right.

True, H.R. 354 does allow "independent creation" of databases in Section 1403(c) and Section 1403(a) exempts nonprofit educational, scientific, and research uses from liability for causing harm to "potential" markets and for certain other reasonable uses.¹⁸⁸ But such defenses in the proposed database regime are no more curative of these constitutional flaws than they would be in the copyright regime, *for the reason that no one can constitutionally oblige all persons not to use facts or ideas that*

fred C. Yen, *A First Amendment Perspective on the Idea/Expression Dichotomy and Copyright in a Work's "Total Concept and Feel,"* 38 EMORY L.J. 393 (1989).

183. See 17 U.S.C. § 102(b) (1994).

184. 499 U.S. 340 (1991).

185. 17 U.S.C. § 107(4) (1994). See also *Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984); PATRY, *supra* note 30, at 205-10.

186. 154 F.2d 464 (2d. Cir. 1946), *cert. denied*, 330 U.S. 851 (1947). See generally PAUL GOLDSTEIN, COPYRIGHT §§ 7.1.1.2, 7.4.1.2 (1998) (discussing infringement in terms of protection of authors' market interest).

187. H.R. 354, 106th Cong. § 1402 (1999).

188. See *id.* §§ 1403(c), 1403(a)(1) (1999) (listing "certain nonprofit educational, scientific, or research uses"); see also *id.* § 1403(a)(2) (listing "Additional Reasonable Uses").

have been made available to the public. Facts and ideas that the copyright law must leave to unrestricted public use cannot constitutionally be withdrawn from public use under the First Amendment by a database law that protects against extraction and use on both primary and derivative markets.¹⁸⁹

In this connection, one should recall that the copyright law, unlike the patent law, does not protect against *use* as such of even the protected expression, as the Supreme Court established in *Baker v. Selden*.¹⁹⁰ The protection of noncopyrightable data and facts against *use* on both primary and secondary markets thus impermissibly disrupts the balance established in the federal copyright and patent laws, which implement the constitutional Enabling Clause.¹⁹¹ Notwithstanding the public's right to use facts and ideas under the First Amendment and notwithstanding the constraints limiting Congressional action under the constitutional Enabling Clause, H.R. 354 institutes copyright-like protection for the use of noncopyrightable matter, creates a *de facto* derivative work right in noncopyrightable compilations, and prohibits transformative uses—however pro-competitive in nature—that harm this reserved or derivative market on the “potential harm” test.

No invocation of unfair competition law can disguise the fact that a “harm to actual or potential markets” test that does not focus on unfair or improper conduct expresses the language of exclusive property rights, which is exactly the function that Section 107(4) performs in the Copyright Law.¹⁹² The fact that it is often physically or economically impracticable to regenerate scientific and research data from scratch only enhances the potential restraints on free speech under H.R. 354 as it stands, by risk-

189. See Letter from Hamilton, *supra* note 139. See also Malla Pollack, *The Right to Know? Delimiting Database Protection at the Juncture of the Commerce Clause, the Intellectual Property Clause, and the First Amendment*, 17 CARDOZO ARTS & ENT. L. J. 47, 67-74 (1999).

190. See *Baker v. Selden*, 101 U.S. 99 (1879); Brown, *supra* note 28. Only those uses specified in the Act are protected. 17 U.S.C. §§ 106, 106A (1994).

191. See U.S. CONST. art. I, § 8, cl. 8; 17 U.S.C. § 301 (1994) (preemption); *Bonito Boats Inc. v. Thunder Craft Boats Inc.*, 489 U.S. 141 (1989); *Compco Corp. v. Day-Brite Lighting Inc.*, 376 U.S. 234 (1964); *Sears, Roebuck & Co. v. Stiffel Co.*, 376 U.S. 225 (1964).

192. 17 U.S.C. § 107(4) (1994); *supra* note 185; see also Letter from Professor Harvey Perlman, Co-Reporter, RESTATEMENT (THIRD) OF UNFAIR COMPETITION LAW (1994), to Senator Orrin Hatch, Chairman of the Senate Committee on the Judiciary (Sept. 8, 1998) (on file with authors) (explaining why neither H.R. 354 nor its predecessors are, in fact, unfair competition laws) [hereinafter Perlman letter].

ing the withdrawal of facts and data as such from the public domain.¹⁹³ The federal appellate courts have consistently declared that avoiding the costs of regenerating known facts and ideas constitutes a basic economic premise underlying the constraints on intellectual property protection deriving from both the First Amendment and the Enabling Clause.¹⁹⁴

The broad definitions of both "collection of information"¹⁹⁵ and "information"¹⁹⁶ in H.R. 354 aggravate these constitutional infirmities by drawing "works of authorship" into the realm of a competing and overlapping intellectual property right, and also by casting legal doubts upon the future ability of third parties to make untrammelled use of public domain matter. Anytime someone would use data, including historical data, that are made available to the public contained in a "collection of information" protected by the proposed law, that user would be exposed to claims that he or she will have harmed the database originator's actual or potential markets if that producer had also used the same or similar data. This broad risk of liability cannot fail to have a chilling effect on the use of known facts and noncopyrightable databases in both the commercial and non-commercial spheres; and it is of little consolation to researchers and educators that they must fear only harm they might cause to actual markets, rather than to potential markets, as well.¹⁹⁷

C. Erring on the Side of Caution

In contrast, a true unfair competition approach would attach liability only when the third party harmed the database maker's actual or potential

193. See *supra* text accompanying notes 120-21.

194. See, e.g., *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991); *Harper & Row v. Nation Enter.*, 471 U.S. 539 (1985). As the Department of Justice's Office of the Legal Counsel recently affirmed, to the extent that the proposed legislation would prohibit extractions or uses of substantial portions of factual compilations by direct competitors, it is much more likely to be held constitutional than if it would prohibit extractions or uses by potential consumers for noncommercial purposes. By contrast, if the provision were construed to provide protection against uses by potential consumers, and not simply direct competitors, it would appear to be of almost limitless scope and therefore to raise constitutional concerns that would appear insurmountable.

DOJ Memorandum, *supra* note 139, at 8.

195. See H.R. 354, 106th Cong. §1401(1) (1999).

196. See *id.*, §1401(2).

197. See *id.*, §1403(a)(1).

market *by improper, unfair, or dishonest means*.¹⁹⁸ Such an approach would not inhibit competitors who “harm” the market by honest and innovative means, and it would not impede true transformative uses that promote competition and the public interest in science and education.¹⁹⁹

The “actual or potential markets” test is thus so broad that it would hinder fair competition simply because every successful competitor harms a prior entrant’s market by definition and because would-be competitors would never know in advance when the use or extraction of protected data may turn out to cause harm to some unknown potential market. In this and other respects, the “harm to markets” test actually cloaks a reserved market formula, in the manner of the exclusive rights to reproduce and to prepare derivative works granted by the Copyright Law.²⁰⁰ Use of this formula in the database context invites other industries to apply for similar protection against harm to their actual or potential markets; and the cumulative anti-competitive effects of recognizing such special-interest protectionist pleas could seriously undermine the ability of the United States to compete in an integrated global marketplace.

United States intellectual property law and policy have traditionally mandated the unfettered use of noncopyrightable facts and of subpatentable ideas, and have favored unbridled competition with respect to the products of mere investment.²⁰¹ We shall undoubtedly experience suboptimal investment in the production of databases if Congress fails to protect publishers against certain forms of piratical conduct that threaten to deprive them of the fruits of their investment.²⁰² But if we combat this risk of market failure by enforcing strong monopolies in collections of data, we may end up balkanizing the information economy by recreating the medieval economic quandary in which products could not flow across countries or continents because too many feudal monopolists demanded payments every few miles down the road.²⁰³

If we discourage follow-on innovation and public good uses of the very databases whose development statutory legal protection is supposed

198. Compare RESTATEMENT (THIRD) OF UNFAIR COMPETITION LAW § 38 (1996) (rejecting general misappropriation doctrine) and Perlman letter, *supra* note 192 with Gordon, *supra* note 7 (proposing tort of malcompetitive copying).

199. Cf. Reichman & Samuelson, *supra* note 7, at 137-45.

200. See 17 U.S.C. §§ 106(1)-(2) (1994).

201. See Reichman, *Charting*, *supra* note 3, at 485-96 (“Negative Economic Premises Underlying the Dominant Legal Paradigms”); *supra* note 191.

202. See, e.g., Tyson & Sherry, *supra* note 72; Hunsucker, *supra* note 7.

203. Cf. Heller & Eisenberg, *supra* note 130; Heller, *supra* note 179.

to stimulate, the end result may be bad for the database industry as a whole²⁰⁴ and devastating for our whole scientific and technical innovation system, which depends on the relatively unrestricted flow of factual data. Instead, we need a regime that loosely preserves a balanced relationship between public and private interests, which courts can develop gradually in response to the empirical conditions of the evolving information economy.²⁰⁵

In the Information Age, as in the Industrial Revolution, we should continue to believe that competition is the lifeblood of commerce, and we should accordingly structure all legal entitlements so as to produce a high degree of competition and maximum dissemination of data and information. If we err on the side of caution and underprotect the building blocks of knowledge, we can always adjust the level of protection upwards later on, in the face of compelling empirical evidence of real economic harm.

But the opposite is not true. Acquired rights and legislatively enacted monopolies cannot easily be eradicated. The wrong decisions today could lessen the vitality of our research enterprise, weaken the national system of innovation, and compromise our future technological superiority, which all depend on maintaining an appropriate balance between upstream and downstream uses of data and factual information.

204. See Testimony of CCIA et al., *supra* note 101.

205. See also Pamela Samuelson (unpublished and untitled essay, manuscript on file with authors) (explaining the need for a national information policy detached from intellectual property policies); Litman, *supra* note 16; Reichman & Franklin, *supra* note 9, at 968-70.

COMMENT

**ELECTRONIC COMMERCE, HACKERS, AND THE
SEARCH FOR LEGITIMACY: A REGULATORY
PROPOSAL**

By Michael Lee, Sean Pak, Tae Kim, David Lee, Aaron Schapiro, and Tamer Francis[†]

ABSTRACT

The escalation of electronic commerce offers a wealth of opportunity for businesses. This technological revolution may be undermined by consumers wary of the increased threat of online invasions of privacy through hacking. The authors detail the various types of security infiltrations—both beneficial and detrimental—that hackers can perpetrate. After examining the current state of federal laws governing hacking, namely the Consumer Fraud and Abuse Law of 1984, the authors posit their recommendations for a realistic regulatory proposal based on an understanding of current technological capabilities.

TABLE OF CONTENTS

| | | |
|-----|---|-----|
| I. | THE THREAT TO ELECTRONIC COMMERCE..... | 844 |
| II. | THE STRUGGLE FOR CODE | 845 |
| | A. The Arsenal..... | 846 |
| | B. Methods and Tools of Attack..... | 847 |
| | 1. <i>Eavesdropping and Packet Sniffing</i> | 847 |
| | 2. <i>Snooping and Downloading</i> | 848 |
| | 3. <i>Tampering or Data Diddling</i> | 848 |
| | 4. <i>Spoofing</i> | 849 |
| | 5. <i>Jamming or Flooding</i> | 849 |
| | 6. <i>Injecting Malicious Code</i> | 849 |
| | 7. <i>Cracking Passwords, Codes, and Keys</i> | 850 |
| | 8. <i>Exploiting Flaws in Design, Implementation or Operation</i> | 850 |
| | C. Countermeasures..... | 850 |
| | 1. <i>Encryption (Secrecy)</i> | 850 |
| | 2. <i>Authentication (Password Systems)</i> | 852 |
| | 3. <i>Access Control and Monitoring (Firewalls)</i> | 852 |

© 1999 Michael Lee, Sean Pak, Tae Kim, David Lee, Aaron Schapiro, and Tamer Francis.

† Michael Lee, Sean Pak, Tae Kim, Aaron Schapiro, and Tamer Francis are third-year law students at Harvard. David Lee is an associate with the law firm of Shearman & Sterling.

This comment was the second-place winner of the 1998 *Berkeley Technology Law Journal* Comment Competition.

| | | |
|------|---|-----|
| 4. | <i>Auditing (Logging) and Intrusion Detection</i> | 853 |
| 5. | <i>Virus Scanners and Disinfectors</i> | 854 |
| 6. | <i>Backup</i> | 854 |
| 7. | <i>Secure Design, Implementation, and Operation</i> | 854 |
| D. | Related Activities: Cracking, Phreaking, Social Engineering | 854 |
| 1. | <i>Cracking</i> | 855 |
| 2. | <i>Phreaking</i> | 857 |
| 3. | <i>Social Engineering</i> | 858 |
| E. | Password Systems: An Arms Race of the Past | 859 |
| F. | Java-Based Security Holes and Safeguards: The Arms Race of the Future | 861 |
| G. | Implications for Regulating Hackers | 864 |
| III. | THE STRUGGLE FOR NORMS | 865 |
| IV. | CURRENT AND PROPOSED LEGAL REGIMES | 868 |
| A. | Hacking as Crime: The Computer Fraud and Abuse Law of 1984 | 868 |
| 1. | <i>The Text of the CFAA</i> | 869 |
| 2. | <i>Access Denied—Access as Crime</i> | 869 |
| 3. | <i>A Critical Evaluation</i> | 872 |
| B. | Hacking as Tort: The Internet Service Provider (“ISP”) Solution | 874 |
| 1. | <i>The Decisive Advantages of a Negligence Regime</i> | 874 |
| 2. | <i>A Critical Evaluation</i> | 877 |
| V. | A HEURISTIC MODEL FOR REFORM | 879 |
| A. | Consider All Relevant Modalities of Regulation | 879 |
| B. | Analyze the Political Consequences of Inducing Changes In Code | 881 |
| VI. | A PROPOSAL FOR OPTIMAL REGULATION | 882 |
| A. | Advantage One: Promotes Self-Regulation Through Market Forces | 883 |
| B. | Advantage Two: Facilitates Democratization of Architectural Developments | 885 |

You bring me a select group of 10 hackers and within 90 days, I'll bring this country to its knees.¹

— Jim Settle, Former Director, FBI Computer Crime Squad

Yes, I am a criminal. My crime is that of curiosity. My crime is that of judging people by what they say and think, not what they look like. My crime is that of outsmarting you, something that you will never forgive me for.²

— the “Mentor”

According to many experts in academia and industry, cyberspace will one day replace real space as the preferred medium for conducting busi-

1. Chris O'Malley, *Information Warriors of the 609th (The Air Force's 609th Information Warfare Squadron)*, POPULAR SCIENCE, July 1997, at 74.

2. The Mentor, *The Mentor's Last Words* (visited Apr. 16, 1999) <<http://insane.bloodline.com/mentor.html>>.

ness.³ Indeed, 1998 was a record year for electronic commerce,⁴ with more than nine billion dollars of retail online sales.⁵ The sudden increase in the volume of electronic commerce has prompted many experts to adjust upward their forecasts for the growth of electronic commerce.⁶

Underlying this optimistic picture for electronic commerce, however, is the basic assumption that consumers and companies will be able to establish what Peter Denning has termed "trust" in the exchange transaction between buyers and sellers in cyberspace.⁷ Despite the theoretical advantages of conducting commerce in cyberspace⁸ and the exponential expansion of the Internet,⁹ many consumers continue to have little confidence in

3. For example, the theory of friction-free markets once posited that Bertrand competition would necessitate pure price competition on the Internet, such that Internet markets would have lower prices than real space markets. See Joseph Bailey & Erik Brynjolfsson, In Search of "Friction-Free Markets": An Exploratory Analysis of Prices for Books, CDs and Software Sold on the Internet, at 3-5 (1998) (unpublished manuscript) (on file with authors).

4. Although Internet-based commerce is the most visible form of electronic commerce, the former is clearly a subset of the latter. As used in this paper, the term "electronic commerce" encompasses all commercial transactions involving the exchange of "bits" as opposed to "atoms."

5. See Commerce Department to Measure Online Sales' Impact (visited Feb. 5, 1999) <http://www.internetnews.com/ec-news/article/0,1087,archive_4_65111,00.html>. There exists a wide variation in estimates of online shopping due to differences in terminology and methodology. See Maryann Jones Thompson, *Spotlight: Why E-commerce Forecasters Don't Get It "Right,"* THE INDUSTRY STANDARD, Mar. 1, 1999, available at <<http://www.thestandard.com/metrics/display/0,1283,850,00.html>>.

6. See Thompson, *supra* note 5.

7. See Peter J. Denning, *Electronic Commerce, in* INTERNET BESIEGED: COUNTERING CYBERSPACE SCOFFLAWS 385-86 (Dorothy E. Denning and Peter J. Denning eds., 1997). Denning argues that cyberspace "trust" would be less difficult to establish with reliable authentication technology, i.e., that current cyberspace code precludes a social norm of "trust." According to Denning, "If human coordination, rather than information exchange, had been at the center of attention of protocol designers, it would be exceedingly difficult today to spoof an e-mail or Internet address or to forge a signature on a document." *Id.* Indeed, building trust online is the focus of many Internet-related companies and consultancies. See Maryann Jones Thompson, *E-commerce Spotlight: Building Trust Online*, THE INDUSTRY STANDARD, Jan. 25, 1999, available at <<http://www.thestandard.com/metrics/display/0,1283,829,00.html>>.

8. Critics have argued that electronic commerce on the Internet reduces overall transaction costs (e.g. search costs, negotiation, and delivery costs) and facilitates connectivity so as to eliminate considerations of real space time and distance. See, e.g., Denning, *supra*, note 7, at 377-78.

9. According to a 1997 Robertson & Co. report, the total number of U.S. Internet users is expected to reach 102 million by the year 2000. See ComputerWorld, *Commerce*

the ability of sellers to deliver goods and services without compromising the security of sensitive information.¹⁰ In fact, retail purchasers on the Internet still represent only a tiny fraction of all consumer spending.¹¹ The most visible agents of distrust have been individuals loosely described as "hackers."¹² Under constant media and governmental scrutiny, hackers have come to occupy a prominent and often mythical role in the popular discourse on electronic commerce.¹³

Surprisingly, however, academic discourse has failed to adequately address the challenges and opportunities posed by hackers for the regulation of cyberspace and electronic commerce. Some critics¹⁴ of regulation have simply cited hackers to further the ambitious claim that attempts to regulate the Internet and its activities are "futile" in general.¹⁵ Others have taken the contrary position that cyberspace actually facilitates effective regulation and that technological solutions will ultimately eliminate the threat to electronic commerce posed by hackers.¹⁶ Somewhere between these opposite ends of the spectrum, Lawrence Lessig has posited that hackers pose little threat or relevant disorder to a regulatory scheme in

by *Numbers* (visited Apr. 9, 1999) <<http://www.computerworld.com/home/Emmerce.nsf/All/pop>>.

10. In a 1999 national survey conducted by Netzero, more than 53 percent of the respondents cite "privacy and security" as their biggest concerns regarding online shopping. See Beth Cox, *Security, Privacy Remain Top Consumer Concerns*, (visited Apr. 9, 1999) <http://www.internetnews.com/ec-news/article/0,1087,4_95031,00.html>.

11. See Greta Mittner, *E-commerce Companies Rejoice*, RED HERRING, Jan. 4, 1999, available at <<http://www.redherring.com/insider/1999/0104/news-shopping.html>>.

12. See discussion *infra* Part I.B.

13. See discussion *infra* Part I.

14. These critics can be characterized either as optimists or pessimists, depending on how one views the broader implications of the advent of electronic commerce. See discussion *infra* Part V.C.

15. See, e.g., David G. Post, *Anarchy, State and the Internet: An Essay on Law-Making in Cyberspace*, 1995 J. ONLINE L. ART. 3 (visited Jan. 20, 1998), available at <<http://www.law.cornell.edu/jol/post.html>>. See also David R. Johnson & David Post, *Law and Borders—The Rise of Law in Cyberspace*, 48 STAN. L. REV. 1367 (1996).

16. For example, Jeffrey Schiller, a computer security expert at M.I.T., claims that encryption technology such as PGP ("Pretty Good Privacy") can provide security against most hacking attacks and that "at this early stage, the insecurity of the Internet is primarily a result of human error and lack of user security education initiatives." Catherine Therese Clarke, *From CrimINet to Cyber-Perp: Toward an Inclusive Approach to Policing the Evolving Criminal Mens Rea on the Internet*, 75 OR. L. REV. 191, 231-32 (1996) (internal quotations omitted).

which the “code” or the architecture of the Internet permits *ex ante* constraints on the vast majority of the inhabitants of cyberspace.¹⁷

Although several distinct models for analyzing the regulation of cyberspace and electronic commerce have emerged with the development of the academic debate, these models, along with current legislation, share an undue emphasis on and reverence for the unique implications of Internet technology. All the models described above gauge the threat posed by hackers—indeed, their very relevance in the regulation debate—solely in terms of the technology or code by which hackers operate. Unfortunately, this (mis)understanding of new technology has precluded analyses of issues equally relevant to an informed discussion on the optimal regulation of cyberspace for the purposes of promoting electronic commerce. Issues meriting further study include: (1) the precise nature of the threat to electronic commerce posed by hackers and their tools, (2) the failure of current and proposed legislation to regulate hackers, and, finally, (3) the broader political nature of cyberspace code and its implications for regulating hackers.

17. According to Lessig:

We live life subject to the code [in cyberspace], as we live life subject to nature. Just as we do not choose whether to see through a wall or not, we don't choose whether to enter America Online without giving our password. Superman might choose whether to see through a wall; and hackers might be able to choose whether to enter AOL with a password. But we are neither supermen or hackers (if such a distinction exists). We live life subject to the constraints of the code; however (and by whomever) these constraints have been set.

Lawrence Lessig, *The Constitution of Code: Limitations on Choice-Based Critiques of Cyberspace Regulation*, 5 *COMMLAW CONSPECTUS* 181, 184 (1997) [hereinafter *Constitution of Code*]. In defense of his claim that code-based solutions for regulating cyberspace are effective despite hacking, Lessig has further stated:

But from the fact that ‘hackers could break any security system,’ it no more follows that security systems are irrelevant than it follows from the fact that ‘a locksmith can pick any lock’ that locks are irrelevant. Locks, like security systems on computers, will be quite effective, even if there are norm-oblivious sorts who can break them.

Lawrence Lessig, *Reading the Constitution in Cyberspace*, 45 *EMORY L.J.* 869, 896 n.80 (1996) [hereinafter *Constitution in Cyberspace*]. Admittedly, Lessig does not claim that hackers do not pose any threat to electronic commerce. Rather, his discussion of hackers is limited to their effect on the long-term architectural development of the Internet, apart from their role in electronic commerce.

I. THE THREAT TO ELECTRONIC COMMERCE

Hacking via the Internet is currently a significant problem, with trends indicating cause for alarm. The business losses from such intrusions can be massive: MCI lost over fifty million dollars when hackers downloaded more than 50,000 credit card numbers,¹⁸ and Citibank lost ten million dollars when its computer network was compromised by a crime group in Russia.¹⁹ The service, repair, and restoration costs from such intrusions are also extensive. For example, in *United States v. Morris*,²⁰ the labor costs to eradicate a computer virus and monitor the computer systems' recovery was estimated at up to \$186 million.²¹

Although these highly publicized cases illustrate the enormous power that a single hacker or a group of hackers may yield, the economic threat posed by hackers is not confined to a handful of Fortune 500 companies. A recent survey by Ernst & Young found that of 1,290 businesses, nearly half had been the victims of information security breaches in the past two years,²² and at least twenty of these companies had suffered losses exceeding one million dollars.²³ According to a Senate report, major banks and corporations lost \$800 million due to hacker intrusions in 1995 alone.²⁴ Moreover, businesses are continually under attack from multiple sources. For instance, Rockwell International, Inc., claims that hackers attempt to break into the company's computers via the Internet on a "regular basis."²⁵

Yet the problem is almost certainly much more extensive than suggested by the available statistics, as many businesses are reluctant to admit that their computers have been successfully attacked by hackers.²⁶ According to William J. Cook, author of the Justice Department's manual on computer prosecution, "[O]rganizations often swallow losses quietly rather than notifying the authorities and advertising their vulnerability to

18. See David L. Gripman, *The Doors are Locked but the Thieves and Vandals are Still Getting In: A Proposal in Tort to Alleviate Corporate America's Cyber-Crime Problem*, 16 J. MARSHALL J. COMPUTER & INFO. L. 167, 169-70 (1997).

19. See Marc D. Goodman, *Why the Police Don't Care About Computer Crime*, 10 HARV. J.L. & TECH. 465, 472 (1997).

20. 928 F.2d 504, 505-07 (2d Cir. 1991).

21. See Gripman, *supra* note 18, at 171.

22. See Marc S. Friedman & Kristin Bissinger, *Infojacking: Crimes on the Information Superhighway*, 9 No. 5 J. PROPRIETARY RTS. 2, 7 (1997).

23. See Goodman, *supra* note 19, at 472.

24. See Friedman & Bissinger, *supra* note 22, at 7.

25. *Id.*

26. See *id.* at 2.

shareholders and clients.”²⁷ Federal law enforcement officers estimate that over ten billion dollars worth of data is stolen in the United States annually,²⁸ and that reports of computer intrusion from government agencies and private businesses jump seventy percent every year.²⁹ According to Dennis Hughes, the FBI’s senior expert on computer crime, “[T]he hackers are driving us nuts. Everyone is getting hacked into. It’s out of control.”³⁰

Careful review of the empirical evidence suggests that hackers pose a significant threat to the future of electronic commerce in two significant ways. First, they *directly* endanger electronic commerce by increasing the risk that private and financial information transmitted over the Internet will be intercepted and used for illegal purposes. Second, they *indirectly* stifle the growth of electronic commerce by undermining the public’s confidence in the safety of conducting financial online transactions. This indirect effect stems largely from the way that consumers, as well as policy makers, perceive hackers. Because hackers are typically characterized as “super-criminals” with extraordinary powers and malicious intent, many consumers may still be afraid to buy and sell goods and services over the Internet, even with adequate safeguards. Thus, the full extent of the economic threat posed by hackers can only be understood by analyzing the phenomenon of hacking from both technological (i.e., code-based) and sociological (i.e., norms-based) perspectives.

II. THE STRUGGLE FOR CODE

The history of the Internet and computer networks in general may be viewed as a story of a continuing arms race between those who seek to erect barriers of protection and those who seek to circumvent these barriers. This story pits governmental organizations, law enforcement officials, and computer professionals against a diverse and ever-expanding group of “[h]ackers, crackers, snoops, spoofers, spammers, scammers, shamblers, jammers, intruders, thieves, purloiners, conspirators, vandals, Trojan horse dealers, virus launchers, and rogue program purveyors.”³¹ The object of

27. *Id.*

28. *See id.* at 7.

29. *See id.* at 10.

30. Gripman, *supra* note 18, at 173.

31. Dorothy E. Denning & Peter J. Denning, *Preface* to *INTERNET BESIEGED: COUNTERING CYBERSPACE SCOFFLAWS* at vii (Dorothy E. Denning & Peter J. Denning eds., 1997).

the struggle is the power to control the "code" (i.e., the underlying architecture) of the Internet.

A. The Arsenal

In *Cyberspace Attacks and Countermeasures*, Dorothy Denning uses the following table to categorize the known methods, tools of attack, and safeguards for protecting against such attacks. The countermeasures are labeled according to their primary purpose: to prevent attacks (P), to detect their occurrence (D), or to facilitate recovery after an incident (R).³²

| | <i>Encrypt. (secrecy)</i> | <i>Authen. (includ. Crypto)</i> | <i>Access Control, Monitor</i> | <i>Audit, Intrusion Detect.</i> | <i>Virus Scan & Disinf.</i> | <i>Backup</i> | <i>Design, Implem., Operat.</i> |
|-------------------------|-------------------------------|-------------------------------------|------------------------------------|-------------------------------------|-------------------------------------|---------------|-------------------------------------|
| <i>Eavesdropping</i> | P | | | | | | P |
| <i>Snooping Storage</i> | P | | P | D | | | P |
| <i>Snooping Memory</i> | | | P | D | | | P |
| <i>Tampering</i> | | D | P | D | | R | P |
| <i>Spoofing</i> | | PD | | D | | | P |
| <i>Jamming</i> | | | P | D | | | P |
| <i>Injecting Code</i> | | PD | P | D | PD | | P |
| <i>Cracking</i> | | | | | | | P |
| <i>Exploiting Flaws</i> | | | P | D | | | P |

Although the categories used above were not meant to be definitive or comprehensive, they do provide a useful framework for discussing the arsenal of weapons available to hackers and anti-hackers. It is important to note, however, that some of the new Java-based attacks may not fit neatly into any of the above categories.

32. *See id.*

B. Methods and Tools of Attack

When attacking a secure Internet site, a hacker may use one or more of the following methods and tools, serially or in combination, to identify security holes and gain unauthorized access.³³

1. *Eavesdropping and Packet Sniffing*

The Internet, like most networks, is susceptible to eavesdropping (i.e., “the passive interception of network traffic”).³⁴ The preferred method of eavesdropping on the Internet is installing a program packet (commonly referred to as a “packet sniffer”) for monitoring network on a local workstation, an Internet gateway, or router machine, which directs and relays network traffic. According to the Computer Emergency Response Team (“CERT”) Coordination Center, following their initial discovery in 1993, sniffer attacks have allowed hackers to gain unauthorized access to more than 100,000 host machines in the United States alone.³⁵

Once installed (either by a user with legitimate access or by a hacker posing as a legitimate user), packet sniffers can be used to intercept login IDs and passwords, as well as credit card information and private e-mail messages.³⁶ Intercepted login IDs and passwords then can be used to access other secured sites. Empirical evidence indicates that once hackers have logged into a secured system through sniffer attacks, their actions can vary.³⁷ In some cases, the hackers moved on to other systems without damaging or otherwise altering any systems or files. In other cases, however, they engaged in malicious activities, including denial of service, unauthorized possession, compromise of integrity, and destruction of data.

33. For example, after cracking a password, a malicious hacker might pose as a legitimate user, browsing through files to gain confidential and financial information. If root access is acquired, the hacker may also leave a destructive logic bomb or alter login records to conceal his tracks. See Dorothy E. Denning, *Cyberspace Attacks and Countermeasures*, in *INTERNET BESIEGED: COUNTERING CYBERSPACE SCOFFLAWS* 29, 32 (Dorothy E. Denning & Peter J. Denning eds., 1997).

34. *Id.*

35. This estimate is likely to be very conservative. See *id.*

36. Upon installation, a packet sniffer places the /dev/nit interface (a widely installed network utility tool) into “promiscuous mode” and logs the first 128 bytes of all TCP (i.e. Internet) sessions being routed through the compromised host machine. The hacker then periodically accesses the host machine to collect the intercepted information. For a more detailed description of packet sniffers, see E. Eugene Schultz & Thomas A. Longstaff, *Internet Sniffer Attacks*, *PROCEEDINGS OF THE NATIONAL INFORMATION SYSTEMS SECURITY CONFERENCE* 534-541 (Oct. 1995).

37. See *id.* at 141.

2. *Snooping and Downloading*

Other methods for acquiring information without altering it include snooping and downloading data without authorization.³⁸ Rather than monitoring and intercepting network traffic, a hacker can obtain unauthorized access to a secured site by using a cracked password and obtain confidential and financial information by browsing through documents, e-mail messages, password files, and other data stored on disk or memory.³⁹ The hacker will often download data to his or her computer before browsing through them. Snooping and downloading can also be done by insiders, especially ex-employees.

3. *Tampering or Data Diddling*

Instead of just downloading data, a hacker, upon obtaining unauthorized access, can alter or delete files and programs stored on secured systems (commonly referred to as data "tampering" or "diddling").⁴⁰ The potential threat can be especially serious if the hacker is able to obtain root access.⁴¹ An extreme form of tampering attack is the placement of logic bombs, which "detonate" in response to a predefined event.⁴² Upon detonation, a logic bomb may crash the entire system or wipe out entire file systems. Another dangerous form of tampering is replacing system programs with their Trojan horse versions, which "look and feel" like the original program, but execute hidden and often malicious code.⁴³ A popular Trojan horse attack involves a modified login program, which operates normally but has the added function of storing copies of login IDs and passwords in a hidden file. As with snooping, data tampering can also be done by insiders.

4. *Spoofing*

In a "spoofing attack," the hacker deceives the victim into disclosing security or financial information by impersonating other users or computers. This form of attack can be analogized to a con game where "the at-

38. As noted earlier, a hacker can combine packet sniffing and snooping attacks to infiltrate a large number of secured sites.

39. In 1996, two hackers were convicted of downloading 1,700 credit card numbers from a Tower Records computer system that they had infiltrated. *See* Dorothy E. Denning, *supra* note 33, at 33.

40. *See id.* at 33-34.

41. Root access enables the hacker to modify system files and programs and to access personal files of every user on the system.

42. *See* Dorothy E. Denning, *supra* note 33, at 33-34.

43. *See id.*

tacker sets up a false but convincing world around the victim."⁴⁴ Common forms of spoofing attacks include e-mail forgery and looping, where a hacker uses one system as a "springboard" to log into another system in order to conceal his or her identity and location.⁴⁵

5. *Jamming or Flooding*

Otherwise known as "denial-of-service" attacks, jamming or flooding attacks aim to disable or to tie up system resources.⁴⁶ Two common forms of this attack are: (1) consuming all available memory or disk space by flooding the target system with large volumes of e-mail, and (2) tying up network connection resources by sending multiple SYN messages requesting Internet connections. Both methods involve using fake return addresses or anonymous remailers to conceal the identity of the attacker. Jamming or flooding attacks can be used to target commercial websites or individual users. The motivation for these types of attack are often personal.

6. *Injecting Malicious Code*

Injecting malicious code (commonly referred to as "viruses") is another type of hacking attack with potentially devastating effects.⁴⁷ As a general rule, the malicious code is transmitted through an external device (e.g., a floppy disk) or through the network (e.g., e-mail attachments) and is activated when the file or data stream is loaded into memory and executed. Typically, the malicious code is designed to be self-replicating (hence the label "virus"), and consequently it may be difficult to predict or control the extent and scope of the damage.⁴⁸ To avoid detection, virus writers often incorporate encryption or self-modifying code into their viruses. In an interesting twist, cryptoviruses, which encrypt rather than destroy the victim's data, have been employed in Britain for extortion purposes.⁴⁹

44. Edward W. Felten et al., *Web Spoofing: An Internet Con Game* (last modified Feb. 1997) <<http://www.cs.princeton.edu/sip/pub/spoofing.html>>.

45. See Dorothy E. Denning, *supra* note 33, at 35.

46. See *id.* at 36.

47. See *id.* at 37-38.

48. A highly publicized example is the Internet Worm program released by Robert Morris. For a detailed account of the Worm program, see KATIE HAFNER & JOHN MARKOFF, *CYBERPUNK: OUTLAWS AND HACKERS ON THE COMPUTER FRONTIER* 280-81 (1991).

49. See Michael McCormack, *Europe Hit by Cryptoviral Extortion*, *COMPUTER FRAUD & SECURITY BULLETIN*, June 1, 1996, at 3.

7. *Cracking Passwords, Codes, and Keys*

Systems that employ password security schemes or encryption algorithms are susceptible to attacks aimed at guessing or finding (commonly referred to as "cracking") a valid password or encryption key.⁵⁰

8. *Exploiting Flaws in Design, Implementation or Operation*

In addition to obtaining passwords by any of the above methods, hackers can also gain unauthorized access by exploiting undetected security flaws in the design, implementation, or operation of secured systems.⁵¹ These security flaws can arise for a number of reasons, including "software bugs, lack of attention to security, and poor configuration."⁵² New operating systems or architectures, such as Java, are especially susceptible to these types of attack.⁵³ Although many security holes are eventually detected and corrected, new ones inevitably arise, sometimes in the new code designed to fix existing flaws.⁵⁴

C. Countermeasures

As with the various forms of hacking attacks and tools, the following categories of countermeasures are interrelated in that the effective operation of a countermeasure may ultimately depend on the success of other related countermeasures.⁵⁵

1. *Encryption (Secrecy)*

Cryptography, defined as the science of using mathematical algorithms to disguise messages and information, is a powerful tool for protecting against various forms of hacking attacks. When used for purposes of secrecy, cryptographic algorithms can serve as effective countermeasures

50. "Cracking" a password or encryption key (i.e., finding or guessing) should be distinguished from "cracking" a software application (i.e., disabling protection features). See A. Michael Froomkin, *The Metaphor is the Key: Cryptography, the Clipper Chip, and the Constitution*, 143 U. PA. L. REV. 709 (1995).

51. A real-life example is the Network File Service ("NFS") and sendmail programs for the UNIX operating system, both of which originally contained bugs allowing regular users (and hackers posing as users) to obtain root access. See Dorothy E. Denning, *supra* note 33, at 38-39.

52. *Id.* at 38.

53. See discussion *infra* Part II.F.

54. See Dorothy E. Denning, *supra* note 33, at 39.

55. See *id.* at 41.

against eavesdropping and snooping.⁵⁶ Encryption, the subset of cryptography dealing with achieving and maintaining secrecy, involves applying a scrambling function to a given set of data so that only those who possess the right “key” can restore (or “decrypt”) the encrypted data to its original (“cleartext”) form. The strength of an encryption system is usually measured by the amount of effort (in terms of computing time) that would be required to “crack” it (i.e., to derive the original data from its encrypted form) by an outsider who knows the algorithm but not the key (or keys) used.

Two of the most popular encryption schemes are the Data Encryption Standard (“DES”) promulgated by the National Bureau of Standards and the RSA system named after its inventors Ronald Rivest, Adi Shamir, and Leonard Adleman. The DES system is a single-key (“symmetric”) system in which a common secret key is used to both encrypt and decrypt data.⁵⁷ The RSA system, by contrast, is a dual-key (“asymmetric” or “public key”) system in which one key is used to encrypt and a second key is used to decrypt.⁵⁸

The effectiveness of both the DES and the RSA systems has been challenged by critics in recent years. Theoretically, there are three ways to crack an encryption system:⁵⁹ (1) hackers can steal the key or suborn a key-holder; (2) hackers can hope to find a mathematical weakness in the cryptographic algorithm; or (3) hackers can use a “brute-force” method of trying all possible keys until the message is decrypted. While all cryptographic systems are susceptible to the first attack, it appears that current implementations of the DES and RSA systems are also vulnerable to attacks exploiting mathematical weaknesses and those utilizing brute-force methods.⁶⁰ As a result, there is a growing demand within the academic and

56. Cryptographic algorithms can be used for two distinct purposes: secrecy and authenticity. The term “encryption” is generally used to refer to cryptographic systems used only for secrecy. *See id.*

57. Typically, an encryption DES system is implemented by requiring a different session key for each communication and providing a different long-term key used for authenticating the user and for distributing session keys.

58. As noted in the following subsection on authentication, public key systems can be used for authentication as well as encryption purposes.

59. *See* Froomkin, *supra* note 50, at 752.

60. In 1996, a 130-digit RSA key was cracked. RSA Laboratories recommends that keys be at least 230 digits (or more than 768 bits). In June 1997, a 56-bit DEC key was broken after four months of trial and error. According to cryptography experts, the DES algorithm is nearing the end of its useful lifetime. *See* Dorothy E. Denning, *Encryption*

business communities to strengthen existing encryption systems by using longer keys or to adopt alternative systems that are inherently more difficult to crack.⁶¹ In the past the government has resisted these demands for change on the grounds of law enforcement and national security.⁶²

Recently, the government offered a compromise solution based on key-escrow systems, which enable government agencies to keep a copy of the key needed to decrypt all encrypted communications.⁶³ Key escrow systems, in theory, satisfy both the demand for stronger encryption and the need for governmental monitoring of personal communications. These proposals, however, have been criticized on constitutional and technical grounds and have yet to be approved by Congress.⁶⁴

2. *Authentication (Password Systems)*

In addition to maintaining secrecy, cryptographic algorithms also can be used for authentication purposes (i.e., to validate that the user is actually who he or she claims to be). If implemented properly, cryptographic systems can prevent against tampering, spoofing, and malicious code attacks.

Public-key encryption systems, such as the RSA scheme, are especially useful as authentication tools.⁶⁵ A user can validate his or her identity by encrypting the message with his or her private key. Upon receipt, the receiver will attempt to decrypt the encrypted message by using the sender's public key, which is freely accessible to all.⁶⁶ If the message has been altered or sent by an impostor, the verification will fail.

3. *Access Control and Monitoring (Firewalls)*

As a countermeasure against snooping and tampering, system designers can incorporate various methods and tools for monitoring and control-

Policy and Market Trends, in INTERNET BESIEGED: COUNTERING CYBERSPACE SCOFFLAWS 458 (Dorothy E. Denning & Peter J. Denning eds., 1997).

61. *See id.* at 457-60.

62. *See* Froomkin, *supra* note 50, at 711.

63. *See id.* at 711-17.

64. *See id.* at 717-51.

65. An example of an RSA-based authentication scheme is the Pretty Good Privacy ("PGP") developed of Phil Zimmerman of MIT. For a more detailed analysis of public-key encryption systems, see Thomas Y.C. Woo and Simon S. Lam, *Authentication for Distributed Systems, in* INTERNET BESIEGED: COUNTERING CYBERSPACE SCOFFLAWS at 319-56 (Dorothy E. Denning & Peter J. Denning eds., 1997).

66. The authenticity of the public key can be guaranteed by a trusted third party (e.g., a certification authority or a member of "a web of trust").

ling access to their secured systems.⁶⁷ For instance, UNIX systems only allow users with root accounts to access certain systems programs and data files. In addition, every program or file on UNIX can be configured with an access control list that specifies which accounts can read write, execute, or search that program or file.

Another example of an access and monitoring system is a firewall, which is placed between an organization's internal network (e.g., Intranet) and the Internet. By using a combination of password, packet filtering and encryption methods, firewalls can be designed to keep out unwanted intruders, exclude undesirable content, and prevent viruses.

4. *Auditing (Logging) and Intrusion Detection*

Most forms of hacking attacks (except for eavesdropping and cracking methods) can be detected by the use of auditing and intrusion detection systems. Auditing systems, which keep records of login activities, can serve as a valuable resource for detecting possible security breaches and for gathering evidence in support of an investigation or prosecution. Intrusion detection systems ("IDS") can provide greater security by enabling real-time detection of intrusion attempts.⁶⁸ IDS systems generally fall into two main categories:

- 1) Anomaly detection systems: Based on the assumption that all intrusive activities are necessarily anomalous, system designers can detect intrusion attempts by comparing current account activities against a "normal activity profile." Commonly used methods of comparison are statistical analysis, predictive pattern generation, and neural networks.
- 2) Misuse detection systems: Similar to virus scanners, misuse detection systems seek to detect intrusion attempts by searching for known attack patterns. The challenge is to distinguish legitimate account activities from known "bad" behavior.

67. For a detailed description of monitoring systems, see Dorothy E. Denning, *supra* note 33, at 45-47.

68. For an in-depth analysis of intrusion detection systems, see Aurobindo Sundaram, *An Introduction to Intrusion Detection* (visited Apr. 16, 1999) <<http://www.cs.purdue.edu/coast/archive/data/author3.html>>.

5. *Virus Scanners and Disinfectors*

Virus scanners are designed to detect the presence of malicious code by looking for signs or patterns of known viruses.⁶⁹ These programs can be configured to scan floppy disks, system memory, or network connections for virus signatures. Once detected, viruses can be removed by disinfectors. Virus scanners and disinfectors are ineffective against newly introduced or custom designed viruses.

6. *Backup*

Because it is difficult to detect and prevent hacking attacks in real time, backing up system data is essential to recovery from accidental and intentional data tampering (e.g., file deletions, virus programs, and logic bombs).⁷⁰ Backup systems, however, cannot prevent the unauthorized downloading and distribution of confidential and financial information.

7. *Secure Design, Implementation, and Operation*

Although no system can be made perfectly secure, all of the hacking attacks discussed above can be countered by making security a top priority in designing, implementing, and operating network systems. Useful tools and methods include good software engineering practices, formal methods, testing, and vulnerability analysis, configuration management, human practices, and user training.⁷¹ Designers and users of secured systems would do well to heed Andy Grove's advice: "Only the paranoid survive."⁷²

D. Related Activities: Cracking, Phreaking, Social Engineering

Most of the so-called "hackers" also engage in a wide range of other activities, including cracking, phreaking, and social engineering. Understanding the tools and methods of these activities is important for two reasons. First, many of the publicized attacks on government and corporate sites have involved a combination of hacking, phreaking, and social engineering tactics. Thus, protecting against hacking attacks alone may not be sufficient to secure an Internet site. Second, from a policy perspective, any regulatory framework designed to control hacking may have an unexpected impact on the ability to control other types of activities. For in-

69. For a discussion of virus scanners and disinfectors, *see* Dorothy E. Denning, *supra* note 33, at 48-49.

70. *See id.* at 49.

71. *See id.* at 49-50.

72. Andy Grove, *ONLY THE PARANOID SURVIVE* (1996).

stance, an effective ban on hacking tools and methods may encourage hackers to employ other means (e.g., cracking, phreaking, or social engineering) in their pursuit of thrills or profit.⁷³

1. *Cracking*

To prevent unauthorized copying and use, software developers have incorporated various protection features into their programs. Some of the more common protection features include:⁷⁴

- 1) Password protection: The user must supply a password before using the program.
- 2) Serial number protection: The user must supply a valid serial number before using the program.
- 3) Use limitation: The user can only use the program a given number of times without paying.
- 4) Time limitation: The user can only use the program for a fixed period of time without paying.
- 5) Disabling some of the functions: The user can only invoke all of the functions of the program upon payment.
- 6) Disk access / token protection: The user must insert a special disk into the floppy drive or attach a special device (i.e., token) to an input/output port before using the program.
- 7) CD access limitation: The user can only use the program if it is stored on a read-only CDROM.
- 8) Any of the above protection features disguised through encryption, "junk" instructions, or self-modifying code.

73. A useful analogy is an underground water reservoir with vertical pipelines. Applying pressure on one of the pipelines will cause water levels to rise in the remaining pipelines. Similarly, applying regulatory pressure on hacking activities may cause incidents of related activities (e.g., cracking, phreaking, and social engineering) to rise.

74. See +ORC ("the old red cracker"), *How to Crack, A Tutorial—Lesson 1* (visited Mar. 13, 1999) <<http://www.geocities.com/Athens/Agora/1948/Crack/howto1.txt>>. The old red cracker, a hacker, has authored one of the many "how-to" manuals on hacking available on the Internet. See *infra* note 76.

All of the above features are designed to make it difficult, if not impossible, for most users to use the protected program without payment or authorization.

“Cracking” is the act of eliminating or suspending one or more protection schemes inside a software application to facilitate unauthorized copying and use.⁷⁵ The most useful tool for a cracker is the “debugger,” a software tool designed to assist software developers in identifying and correcting flaws in a computer program. A debugger allows the user to execute a computer program one instruction or one set of instructions at a time. Another useful tool is the “memory dump analyzer,” which enables the user to examine in detail the memory space of a computer system. By employing these readily available tools in conjunction, a cracker can identify and isolate protection features of a software application and disable them by altering its object code.⁷⁶

Once a protection feature has been disabled, the cracker can automate the process by writing an executable software patch, which can be downloaded and executed by anyone. Cracking patches (commonly referred to as “cracks” or “crackz”) as well as pirated software (commonly referred to as “warez”) can be downloaded from publicly accessible websites.⁷⁷ Valid serial numbers for various software applications are also available on the Web.

By all accounts, cracking activities pose a significant threat to the software industry.⁷⁸ According to a study conducted by the Business Software Alliance (“BSA”) and Software Publishing Association (“SPA”), nearly one of every two new business applications used globally were pi-

75. See *id.* (defining cracking as “understanding, broadly individuating, locating exactly and eliminating or suspending or deferring one or more protection schemes inside a software application you do not possess the source code of”).

76. “How-to” manuals vary in quality and accessibility. An example of a well-written and widely-read manual is the *How to Crack, A Tutorial*, *supra* note 74, written by “the old red cracker.” This manual gives step-by-step instructions on how to crack various types of software applications, including those written for the Windows operating system.

77. See, e.g., Krupt Technologies, *Krupt Warez* (visited Apr. 16, 1999) <<http://www2.ipeg.com/>>.

78. The provision of cracker utilities and serial numbers that are intended to circumvent the copyright protections in software, when used by a direct infringer, may constitute contributory infringement under copyright law. See *Software Publishers Association Policy Statement on Contributory Infringement* (visited Feb. 5, 1999) <<http://www.spa.org/piracy/contrib.htm>>.

rated in 1996, resulting in an estimated \$11.2 billion of lost revenue for the software industry.

2. *Phreaking*

Before the advent of the World Wide Web, hackers communicated with one another via Bulletin Board Systems ("BBS"). BBSs are privately owned and privately operated networks that can be accessed through a dial-up modem connection. As PCs and modems became readily available, various hacking groups began to operate their own underground BBSs, which served as forums for sending messages to other hackers, exchanging information on newly discovered hacking attacks, and sharing pirated software. By the early 1990s, hundreds, if not thousands, of hacker BBSs were in operation across the United States and Europe.

To gain access to a BBS outside the local exchange area, a hacker had to establish a long-distance modem connection, thus giving rise to "phreaking." Phreaking is "a subset of computer hacking and involves hacking of telephone systems to make fraudulent phone calls, or manipulating telephone systems."⁷⁹ By building various devices from off-the-shelf components and by making use of confidential information regarding telephone systems, phreakers are able to make long-distance calls, install calling features like caller-ID or call waiting, and make conference calls—all for free. Some of the more popular phreaking devices include:⁸⁰

- 1) Red Box: Built from a modified Radio Shack tone dialer or a Hallmark greeting card, a Red Box allows the user to make free phone calls by simulating a quarter tone for public telephones.
- 2) Blue Box: Built from the same components as a Red Box, a Blue Box allows the user to convince the telephone system that he or she is actually a telephone operator.
- 3) Black Box: Built from a 10k ohm resistor, a Black Box prevents the phone company equipment from detecting that the user has answered an incoming call. People who call the user's number will not be billed for the call.

79. Jim Christy, *Rome Laboratory Attacks: Prepared Testimony of Jim Christy, Air Force Investigator, before the Senate Governmental Affairs Committee, Permanent Investigations Subcommittee, May 22, 1996*, in *INTERNET BESIEGED: COUNTERING CYBERSPACE SCOFFLAWS* 64 (Dorothy E. Denning & Peter J. Denning eds., 1997).

80. For a more detailed listing of various phreaking devices, see Voyager, *#hack Frequently Asked Questions (FAQ)* (visited Feb. 5, 1999) <ftp://rtfm.mit.edu/pub/usenet-by-group/alt.2600/alt.2600_FAQ>.

In addition to calling long-distance for free, hackers often employ phreaking tools and methods to disguise their calling location, thereby making it difficult, if not impossible, for law enforcement officials to trace suspected hackers back to their origin in real time. It is especially difficult to trace hacking attacks made through multiple paths across multiple systems in multiple countries, as was the case in the highly publicized attacks on the Rome Laboratory of Griffiss Air Force Base during 1994.⁸¹

3. *Social Engineering*

As the least discussed category of hacking-related activities, social engineering may also be the most important. "Social engineering" is a term used within the hacking community for hacking techniques that rely on "weaknesses in wetware [i.e., people] rather than software."⁸² The aim is to trick or deceive people into revealing passwords or other information that may compromise the security of a target system or organization.⁸³ Classic social engineering methods include phoning a "mark" (usually a user or an employee) who has the desired information and posing as a field service technician or a fellow employee with an urgent access problem.⁸⁴

A more sophisticated method is to design and send promotional material (e.g., a fake entry form for a mass-mail sweepstakes) via regular mail to be filled out by the mark or group of marks. One of the entries should be a password for verification (usually associated with a prize) based on the assumption that the mark will enter the same password that he or she uses to gain access to his or her network account. The hacker can then use this password to gain unauthorized access to secured accounts. Other methods include posing as a system operator ("SYSOP") to an unwitting user in an online chat room or going through someone's trash, commonly referred to as "dumpster diving."

Social engineering tactics are often used to complement other methods of hacking and, in some cases, may prove to be the most effective and time efficient way to gain unauthorized access to secured systems.⁸⁵ As a

81. See Christy, *supra* note 79, at 57-65.

82. See bernz, *The Complete Social Engineering FAQ §1.1* (visited Feb. 4, 1999) <<http://members.tripod.com/~bernz/socenfaq.txt>>.

83. See *id.*

84. See *id.*

85. For a real-life account of social engineering, see *The New York Newsday Interview with Ice Man and Maniac: Inside the Underworld of "Hacking,"* N.Y. NEWSDAY, July 22, 1992, at 83.

general rule, it is easier to find lapses in security by the people who use the systems, rather than in the systems themselves.⁸⁶

E. Password Systems: An Arms Race of the Past

In the preceding analysis, various hacking attacks and their countermeasures were described without reference to their related histories. Most security systems, however, are the result of many years of competition between those responsible for maintaining security and those seeking to attack it. Thus, at any given point in time, the design of these systems reflect the ongoing arms race between the system designers and the hackers. For this reason, it is often instructive to trace the history of various security systems in addition to analyzing their present strengths and weaknesses.

One illustrative example is the development of the password security scheme used in the UNIX operating system.⁸⁷ The UNIX system was initially implemented with a password file containing the actual passwords of all the users. This scheme was quickly proven to be “excessively vulnerable” to lapses in security.⁸⁸ The vulnerability stemmed from the fact that there was no way to prevent privileged users from making copies of the password file. Thus, once a hacker had access to a password for privileged user status, he or she had access to all the passwords for the system. In addition, accidental disclosures of the password file jeopardized the security of the entire system. Experiences with earlier remote-access systems indicated that such disclosures occurred with alarming frequency.⁸⁹

In order to remedy these flaws, the UNIX designers added an encryption component to the password system. Before each user password was stored in the password file, it was first encrypted using a modified version of the M-209 cipher scheme used by the U.S. Army during World War II.⁹⁰

86. For instance, the most sophisticated password system can be circumvented by deceiving one of its users to disclose his or her password unwittingly.

87. The following discussion on the history of password security systems is based on Robert Morris & Ken Thompson, *Password Security: A Case History* (visited Jan. 21, 1998) <http://www.securezone.com/Information_Sources/Papers/>.

88. *See id.*

89. *See id.*

90. The problem with the original M-209 scheme was that, with a given key, encrypted messages (or “ciphers”) were trivial to invert. It was much more difficult to reverse engineer the key given the cleartext input and the encrypted messages. Thus, the UNIX designers decided to use the password not as the text to be encrypted, but as the

Theoretically, the encrypted password system was very difficult to penetrate because brute force methods for inverting the encryption algorithm used were prohibitively slow. The system, however, proved to be vulnerable to so-called "key search" attacks. This class of hacking attacks is based on the fact that people tend to choose relatively short passwords that are easy to remember, such as words, names or birth dates. Using this insight, hackers were able to gain unauthorized access to secured systems with encrypted password schemes by comparing the encrypted entries in the password file against a collection of trial passwords that have been encrypted using the same algorithm as the one used by the target system. The success of this method depended on the hacker's ability to decrease the required amount of computing time by carefully choosing the collection of trial passwords. The most successful approaches employed trial passwords derived from a dictionary or list of names.⁹¹

In response to the unexpected success of key search attacks, the UNIX designers adopted the following countermeasures:

- 1) Slower encryption: To increase the amount of computing time required to conduct key search attacks, the M-209 algorithm was replaced with the slower DES encryption algorithm approved by the National Bureau of Standards.
- 2) Less predictable passwords: The password entry program was redesigned to encourage users to adopt longer and more obscure passwords.
- 3) Salted Passwords: To reduce the likelihood of finding a match using a large collection of encrypted password files, the password system was modified to append a randomly generated 12-bit number (called the "salt") to the password typed in by the user before being encrypted and stored in the password file.

With these countermeasures in place, the UNIX operating system is considered to be one of the more secure operating systems on the market. It is important to note that the development of these countermeasures were made possible by the decision (on the part of the UNIX designers) to pub-

key to encrypt a predetermined constant. The encrypted result was then stored in the password file.

91. Some "profitable" entries to include as trial passwords are: (1) the dictionary with the words spelled backwards; (2) a list of first names, last names, and street names; (3) all valid license plate numbers; (4) social security and telephone numbers.

licize the design of the password system and to invite attacks on its security, rather than "playing the customary make-believe game in which weaknesses of the system are not discussed no matter how apparent."⁹²

At the same time, the system is not perfect and remains vulnerable to unanticipated hacking attacks made possible by exploiting a flaw in the implementation or by capitalizing on rapid advances in technology. For instance, brute force methods for inverting the DEC algorithm may become more feasible as the computing power available to the general public continues to increase exponentially.⁹³

F. Java-Based Security Holes and Safeguards: The Arms Race of the Future

As the battle for control of the password system continues, the advent of the World Wide Web and the Java programming language has spawned a new arms race between those seeking to erect barriers of protection and those seeking to circumvent them. Although this new battle is being fought with some of the same weapons used to attack and protect password systems of the past, the unique characteristics of the World Wide Web and the Java language in particular have created new opportunities for hackers.

One of the most powerful features of the Java programming environment is the ability to develop and distribute executable content (commonly referred to as "applets") across heterogeneous platforms.⁹⁴ In effect, Java has transformed the Web from a static collection of mostly textual pages to an interactive and animated world of mini-applications that can be downloaded and executed on any machine on the Internet. With Java applets at their disposal, content providers on the Web now possess unprecedented levels of programming power and expressive potential.

Unfortunately, the very properties that make Java so exciting also make it the greatest threat to Internet security.⁹⁵ If a Java-compatible Web browser is not carefully configured, it can provide a malicious applet with the ability to delete files on the user's personal computer and to send private information over the network surreptitiously.⁹⁶ The solution, how-

92. Morris & Thompson, *supra* note 87, at 5.

93. See discussion *supra* Part II.C.1.

94. For a detailed discussion of the Java programming language and executable content in general, see Joseph A. Bank, *Java Security* (Dec. 8, 1995) <<http://swissnet.ai.mit.edu/~jbank/javapaper/javapaper.html>>.

95. See *id.*

96. See *id.*

ever, is not as simple as completely preventing Java applets from accessing resources on the host machine. Without access to certain resources, Java applets would be of limited value.⁹⁷ For instance, a word processor that cannot save files is useless.⁹⁸ The challenge is to identify the resources required by a particular Java applet and to provide controlled access to those resources without jeopardizing the security of the host machine.

Although the Internet community is only beginning to appreciate fully the dangers posed by uncontrolled Java applets, experts have long recognized the security threats posed by the Java programming environment.⁹⁹ Known Java-based attacks can be organized into five categories:

- 1) Data tampering attacks: Hackers can exploit various implementation flaws in Java to create and distribute malicious applets that modify or delete files and memory locations.¹⁰⁰
- 2) Denial-of-service attacks: Malicious applets can also tie up system resources and crash host machines by consuming processor cycles and misallocating memory resources.¹⁰¹
- 3) Disclosure attacks: Malicious applets can transmit private information stored on the host machine by accessing user files and establishing covert channels with an undisclosed third-party site on the Internet.¹⁰²
- 4) Annoyance attacks: Malicious applets can project offensive video and audio data on the host machine without the user's consent or authorization.¹⁰³
- 5) Web spoofing attacks: Malicious applets can also deceive the user into thinking that he or she is communicating with a trusted site through a secure connection when in fact all submitted information

97. *See id.*

98. *See id.*

99. *See, e.g., id.*; Gary McGraw & Edward Felten, *Understanding the Keys to Java Security—the Sandbox and Authentication*, JAVA WORLD, May 1997, available at <<http://www.javaworld.com/javaworld/jw-05-1997/jw-05-security.html>>; Drew Dean et al., *Java Security: Web Browsers and Beyond*, in INTERNET BESIEGED: COUNTERING CYBERSPACE SCOFFLAWS 241-71 (Dorothy E. Denning & Peter J. Denning eds., 1997).

100. *See Bank, supra* note 94.

101. *See id.*

102. *See id.*

103. *See id.*

is being forwarded to an undisclosed third-party site on the Internet.¹⁰⁴

While Netscape and Microsoft have redesigned their Java-compatible browsers to protect against many of the known attacks, effective countermeasures have not been developed for others, including denial-of-service and Web-spoofing attacks.¹⁰⁵ Although denial-of-service attacks may cause at worst inconvenience for an individual user, Web-spoofing is a dangerous and nearly undetectable security attack that can pose a significant threat to conducting electronic commerce on the Internet.¹⁰⁶

The potential dangers of spoofing attacks on the Web were vividly illustrated by a team of computer science researchers at Princeton University, who created a "shadow copy" of the entire World Wide Web.¹⁰⁷ The researchers placed a spoofing program on one of the servers linking the victim to the rest of the Web (referred in the following discussion as "the hacker's server"). Upon installation, the spoofing program first rewrites all of the URLs on a selected Web page so that they point to the hacker's server rather than to a real server on the Web. If the victim requests a Web page through any of the rewritten URLs, the spoofing program fetches the real page from the Web and modifies the page before forwarding it to the victim. The requested page can be modified to store hidden copies of form entries, passwords, or other information submitted by the victim. To complete the illusion, JavaScript programs can be used to hide all evidence of the spoofing attack from the victim.

One particularly troublesome property of this attack is that it is effective even when the victim believes that he or she is requesting a Web page through a "secure" connection.¹⁰⁸ Because the spoofing program is acting as a hidden intermediary between the victim's computer and the Web, the secure connection is established to the hacker's server, rather than the intended server. Thus, any information transmitted over this connection, whether encrypted or not, is visible to the spoofing program.

As the Princeton researchers noted in their paper, there appears to be no fully satisfactory countermeasure to Web-spoofing attacks, short of disabling the Java feature entirely.¹⁰⁹ Until an effective countermeasure is

104. See Felten et al., *supra* note 44.

105. See Bank, *supra* note 94, at 10.

106. See Felten et al., *supra* note 44.

107. For a detailed description of Web spoofing attacks, see *id.*

108. See *id.*

109. See *id.*

discovered, users conducting secured commercial transactions on the Web are forced to choose between risking undetected disclosures of their financial information and foregoing Java compatibility altogether.

In addition to the security threat posed by Web-spoofing attacks, new and unanticipated Java-based attacks are being discovered on a regular basis. For instance, in February of 1997, anonymous hackers (using the pseudonyms "Major Malfunction" and "Ben Laurie") exposed two new Java-based attacks that "crack a 'secure' client machine wide open."¹¹⁰ The first attack enables a hacker to discover the real identity of any client machine despite the use of precautionary measures such as firewalls, proxies, and SOCKS hosts.¹¹¹ The second and more dangerous attack allows a hacker to scan any TCP/IP port on a client machine.¹¹² Using this attack, a hacker can copy sensitive information transmitted over the compromised port and surreptitiously transmit this information back to his or her machine through a covert channel. In response, Netscape and Microsoft have since released patches to prevent against these types of attacks.¹¹³

The arms race between hackers and Java designers has just begun. No one can predict when the race will end or who will win. What is clear, however, is that Java will play an increasingly larger role in the development of the World Wide Web as "marketspace."

G. Implications for Regulating Hackers

The preceding analysis suggests that problems posed by hackers cannot be solved by technological means alone. History and experience have shown that no system can be made perfectly secure. "Secured systems" employing code-based solutions will always remain vulnerable to unexpected attacks exploiting overlooked flaws in design, implementation, or operation. Moreover, new technologies, such as Java, create new opportunities for users and hackers alike. Often the very properties that make these new technologies exciting and valuable will also give rise to new and unexpected security threats.

110. Although the hackers had employed the newly discovered attacks to hack their way through firewalls in January of 1997, they had decided to give Netscape and Microsoft ample time to address the problem before they publicly disclosed their methods. See Gary McGraw, *Is Your Browser a Blabbermouth? Are Your Ports Being Scanned?*, JAVA WORLD, Mar. 1997, available at <<http://www.javaworld.com/javaworld/jw-03-1997/jw-03-securityholes.html>>.

111. See *id.*

112. See *id.*

113. See *id.*

The insufficiency of code-based solutions is best expressed by Peter and Dorothy Denning in their book *Internet Besieged: Countering Cyberspace Scofflaws*, an anthology of leading experts on Internet security:

We believe that the [problems caused by hacking attacks] are a serious threat to information infrastructures everywhere. Until they are addressed satisfactorily, all the widely touted boons of the Internet—from tele-work to distance education to electronic commerce—will not be realized ... We also believe that the solutions to these problems cannot be achieved solely by technological means. The answer will involve a complex interplay among law, policy, and technology.¹¹⁴

III. THE STRUGGLE FOR NORMS

As hackers battle against system designers for control of the code, they find themselves battling one another for control of the internal norms governing the hacking community. Initially, hackers were a homogeneous and tightly knit community, united by a common desire to learn and a strong code of ethics.¹¹⁵ They viewed themselves as “learners and explorers who want to help rather than cause damage, and who often have very high standards of behavior.”¹¹⁶ In fact, they were generally scornful of those who employed hacking methods and tools for malicious or profit motives.¹¹⁷

The so-called “hacker ethic” included the following principles¹¹⁸:

114. Dorothy E. Denning & Peter Denning, *Preface*, *supra* note 31, at x-xi (emphasis added).

115. See generally Dorothy E. Denning, *Concerning Hackers Who Break into Computer Systems*, at 13 (visited Jan. 23, 1998) <<http://www.cpsr.org>>.

116. *Id.* at 1.

117. Dorothy Denning, in her 1990 survey of the hacking community, stated that, according to all of the hackers she spoke with, malicious hacking was considered morally wrong. They also said that most hackers were not intentionally malicious, and that they were concerned about causing accidental damage. See *id.* at 10.

118. In *A Novice's Guide to Hacking*, the “Mentor,” one of the members of the Legion of Doom hacking group, presents the following set of guidelines for beginning hackers:

Do not intentionally damage any system.

Do not alter any system files other than ones needed to ensure your escape from detection and your future access.

Do not leave your real name (or anyone else's) real name, real handle, or real phone number on any system that you access illegally.

Be careful who you share information with.

- 1) Access to computers—and anything which might teach you something about the way the world works—should be unlimited and total.”¹¹⁹
- 2) “All information should be free.”¹²⁰
- 3) “Thou Shalt Not Destroy.”¹²¹

Surprisingly, the above principles mirror the ethical standards adopted by many computer security professionals. Where the two groups differ is that hackers did not consider the act of breaking into secured systems as inherently unethical.¹²² They believed that hacking was not the same as stealing, but was in fact beneficial because hackers were able to uncover latent design flaws and security deficiencies.¹²³

For the most part, these ethical principles were shared by leaders of the earlier generation of hackers. Most hacker organizations at the time operated by the rule of mutual teaching and learning: If a hacker wanted to learn from other hackers in the group, he or she had to contribute to the knowledge base.¹²⁴ Accordingly, those who accomplished the most and discovered the most creative hacking methods naturally became the leaders. Conversely, those who did not possess the drive to learn were precluded from benefiting from the knowledge of other hackers.

More importantly, these leaders also possessed the means to enforce the “hacker ethic.” Prior to the Internet, hacker organizations existed

Do not leave your real phone number to anyone you don't know.
 Do not hack government computers.
 Don't use codes unless there is no way around it.
 Don't be afraid to be paranoid.
 Watch what you post on boards.
 Don't be afraid to ask questions.
 Finally, you have to actually hack.

A Novice's Guide to Hacking—1989 Edition (visited Apr. 16, 1999) <<http://insane.bloodline.com/mentor.html>>.

119. *Id.* at 5.

120. *Id.* at 5.

121. *Id.* at 10.

122. *See id.* at 10-11.

123. *See id.* at 11. *But see* Eugene H. Spafford, *Some Musings on Ethics and Computer Break-ins* (visited Jan. 19, 1998) <<http://www.cs.purdue.edu>>.

124. *See The New York Newsday Interview with Ice Man and Maniac: Inside the Underworld of "Hacking," supra* note 85. In an interview with a Newsday reporter, Joshua Quitner, a well-known hacker by the pseudonym of “Maniac” stated: “[Hacking] is an organized hobby. You do these things for us and you get a little recognition for it.” *Id.*

largely through closed networks such as bulletin board systems (“BBSs”).¹²⁵ Because these BBSs were privately owned and privately managed, the leaders of hacking organizations had the power to accept only those who pledged to the existing norms and to remove from membership those that violated them. As a result, the original hacking community was a hierarchy based on expertise and knowledge, rather than profit or criminal intent. Moreover, this hierarchy based on accomplishment ensured that those who became the most technologically proficient—and thus had access to the most potentially dangerous knowledge—were those who had gone through the norm-reinforcing process.

Although some critics have questioned whether the so-called “hacker ethic” was the exception rather than the rule within the hacking community,¹²⁶ it is undisputed that such norms did exist and that they did have a profound effect on the way hackers viewed themselves and their activities. It is equally clear, however, that “the hacker ethic is fading fast with the advent of the Internet.”¹²⁷

The Internet drastically changed the internal dynamics of the hacking community in several different but related ways. First of all, the Internet’s open architecture and its increasing accessibility have created huge opportunities for large businesses and individual entrepreneurs alike. As explained in the previous section on electronic commerce, online companies must obtain, transmit, or place commercially valuable information, such as credit card numbers, on the Internet.¹²⁸ With more commercial transactions being conducted on the Internet everyday, the potential profit for malicious hacking activities has grown dramatically. Consequently, the hacking community is increasingly attracting profit-driven and criminally-minded outsiders who do not follow the hacker code of ethics.¹²⁹

Market forces aside, the very architecture of the Internet has made it difficult to maintain the hacker code of ethics. Many hackers have moved away from private and closed networks, such as bulletin board systems, onto the Internet. Whereas BBS-based hacking groups could easily exclude non-members or norm-violators, Internet-based hacking groups do not necessarily have such self-selecting mechanisms. It is now possible for a norm-breaking hacker to distribute his or her knowledge on the Internet

125. See discussion *supra* Part II.D.2.

126. See Benjamin J. Fox, *Hackers and the U.S. Secret Service* (visited Jan. 20, 1998) <<http://www.gse.ucla.edu/iclp/bfox.html>>.

127. *Id.*

128. See discussion *supra* Part I.

129. See Fox, *supra* note 126.

users by simply posting a Web page. Criminals who want to take advantage of such information are finding it increasingly easier to gain access to dangerous hacking tools and methods without ever contributing to the mutual learning process or being subject to the norms that once governed the pre-Internet hacker community.

Although the Internet has certainly resulted in the dilution of norms within in the hacking community, it would be incorrect to say that norms no longer exist. Rather, the hacker community now consists of different sectors with different and often conflicting norms. Thus, the internal struggle for norms is critical to the future of the hacking community and the threat that it may pose to electronic commerce.

IV. CURRENT AND PROPOSED LEGAL REGIMES

The proliferation of various means by which hackers now manipulate the architecture of cyberspace and the growing visibility of hackers willing to misuse these means have not gone unnoticed by lawmakers. Unfortunately, however, existing attempts to regulate malicious hackers have produced dismal results by any standard. Although the sources of such failure are many, perhaps the most debilitating is that the regulatory approach underlying these attempts to control hackers betrays a cursory understanding of the dynamics of a community of social dissidents whose growth and danger have been fueled by the very laws attempting to extinguish it. Indeed, while direct regulation may have proven satisfactory for two centuries of real space regulation, a critical examination of both current laws and existing proposals for reform reveals that an entirely different form of regulation is appropriate for cyberspace.

A. Hacking as Crime: The Computer Fraud and Abuse Law of 1984

Congress has treated computer-related crimes as distinct federal offenses since its enactment of the Computer Fraud and Abuse Law of 1984 ("CFAA"), a seminal piece of legislation embodying the predominant approach to regulating hackers. As mentioned above, the types and sheer number of computer crimes have expanded considerably in a rather short span of time, and the CFAA has since been amended to cover new strains of computer crime facilitated by emerging technologies. Despite numerous amendments, however, the history of the CFAA and attempts to enforce its ever expanding provisions highlight severe inadequacies symptomatic of all current approaches to regulating hackers.

1. *The Text of the CFAA*

The CFAA prohibits “knowingly, and with intent to defraud, access[ing] a protected computer without authorization.”¹³⁰ The text of the statute defines all relevant terms broadly. A “protected computer” is one

exclusively for the use of a financial institution or the United States Government, or, in the case of a computer not exclusively for such use, used by or for a financial institution or the United States Government and the conduct constituting the offense affects the use of the financial institution’s operation or the Government’s operation of such computer.¹³¹

However, a “protected computer” as defined by the CFAA is also one that is used in interstate or foreign commerce or communication.¹³² As a result, any computer with Internet access qualifies as a “protected computer” for purposes of the CFAA.

2. *Access Denied—Access as Crime*

In theory, the CFAA does not prohibit all unauthorized, intentional access to such computers. Unauthorized, intentional access is proscribed only if such access is gained:

- 1) to obtain information relating to national defense or foreign relations. The *mens rea* requirement is that the offender knowingly access a computer without authorization or exceeding authorized access.¹³³
- 2) to obtain information in a financial record of a financial institution or consumer reporting agency, any information from any department or agency of the United States, and information from any protected computer if the conduct involves an interstate or foreign communication.¹³⁴

130. 18 U.S.C. § 1030(a)(4) (1998).

131. 18 U.S.C. § 1030(e)(2) (1998).

132. *See id.*

133. *See* 18 U.S.C. § 1030(a)(1) (1998). To be prosecuted under § 1030(a)(1), the actor must have reason to believe that such information will be used to the injury of the United States or to the advantage of any foreign nation. Further, the section is violated regardless of whether the actor communicates the information to another person or simply retains it. This crime is treated as a felony.

134. 18 U.S.C. § 1030(a)(2) (1998). A “financial record” is defined as “information derived from any record held by a financial institution pertaining to a customer’s relationship with the financial institution.” 18 U.S.C. § 1030(e)(5) (1998). Under this section, ob-

- 3) to manipulate information on any computer that is exclusively for the use of the Government, or in the case of a computer not exclusively for such use, is used by or for the Government, such that the actor's offense adversely affects the use of the computer by or for the Government.¹³⁵
- 4) to access a protected computer, without or in excess of authorization, with the intent to defraud and obtain anything of value, unless the thing obtained consists only of the use of the computer and the value of such use is not more than \$5,000 in any 1-year period.¹³⁶
- 5) to intentionally without authorization access a protected computer, where such access alters or damages program, information, code or command.¹³⁷

taining information of minimal value (\$5,000 or less) results in a misdemeanor, whereas obtaining valuable (more than \$5,000) information or misusing information for financial or commercial gain or to commit a criminal or tortious act constitutes a felony.

135. 18 U.S.C. § 1030 (a)(3) (1998). Section 1030(a)(3) criminalizes electronic trespasses on Federal Government computers. If the computer is not exclusively used by the Government, a violation is found if the trespasser's conduct affects the use of the computer by the Government.

136. 18 U.S.C. § 1030(a)(4) (1998). This section contains a "computer use" exception where the intent to defraud consists only in making use of the computer.

137. 18 U.S.C. § 1030(a)(5) (1998). Section 1030(a)(5) contains three provisions covering both outsider hackers and insiders who cause intentional, reckless or negligent damage. Violating the first two provisions is a felony, violating the third provision is a misdemeanor, with penalties based on the intent and authority of the actor.

The first provision prohibits unauthorized access to a protected computer where the actor knowingly transmits any program, information, code, or command which *intentionally* causes damage, covering both insiders and outsiders. The second provision prohibits unauthorized, intentional access to a protected computer, where such trespass *recklessly* causes damage, covering only outside hackers. The third provision prohibits the same action, but where such trespass causes damage, covering outside hackers. *See* S. Rep. No. 104-357, at 7-8 (1996).

Thus, insiders authorized access to a protected computer face criminal liability only for causing intentional damage, whereas outside hackers who break into a computer can be held liable for intentional, reckless, or negligent damage. This distinction between outsiders and insiders stems from the doctrine of trespass:

To provide otherwise is to openly invite hackers to break into computer systems, safe in the knowledge that no matter how much damage they cause, it is no crime unless that damage was either intentional or reckless. Rather than send such a dangerous message (and deny victims any relief), it is better to ensure that § 1030(a)(5) criminalizes all computer trespass, as well as intentional damage by insiders, albeit at different levels of severity.

- 6) to “knowingly and with intent to defraud,” without authorization, “traffi[c] in any password or similar information through which a computer may be accessed” if “such trafficking affects interstate or foreign commerce; or such computer is used by ... the Government.”¹³⁸
- 7) to extort from any legal entity anything of value, transmitting in interstate or foreign commerce any communication containing any “threat to cause damage to a protected computer.”¹³⁹

Nevertheless, the aggregate effect of these qualifications has been to criminalize all unauthorized, intentional access to protected computers. Of particular significance in the history of the CFAA has been the willingness of Congress to reduce the requisite level of *mens rea* required for prosecution under the statute.¹⁴⁰ Whereas the 1994 amendment classified the requisite *mens rea* as “intentional, knowing, and reckless,” the latest amendment enacted in 1996 eliminated the *mens rea* requirement altogether by imposing strict liability in addition to the requisite *mens rea* as enacted in 1994.¹⁴¹ In short, Congress has criminalized unauthorized access into computer systems, regardless of whether the computer user actually intended to cause damage.¹⁴²

Courts have also interpreted the *mens rea* requirement under the CFAA to facilitate the prosecution of hackers. In *United States v. Mor-*

Id.

The term “damage” is broadly defined to include any impairment to the integrity or availability of data, a program, a system, or information that (A) causes loss aggregating at least \$5,000 in any one-year period to one or more individuals; (B) either modifies or impairs, or potentially modifies or impairs, the medical examination, diagnosis, treatment, or care of one or more individuals; (C) causes physical injury to any person; or (D) threatens public health or safety. *See* S. Rep. No. 104-357, at 8 (1996).

However, it is unclear whether there is a loss if, for example, a virus does not destroy files, but simply overloads the network, thus slowing down processing speed or using up some of a system’s underutilized capacity. What is clear is that this section was added to address the threat posed by hackers. *See* S. Rep. No. 104-357, at 9 (1996) (describing § 1030(a)(5) as a measure that protects computers from hackers).

138. 18 U.S.C. § 1030(a)(6) (1998).

139. 18 U.S.C. § 1030(a)(7) (1998).

140. *See* S. Rep. No. 104-357, at 10-11 (discussing changes in *mens rea* level).

141. *See* S. Rep. No. 104-357, at 9-12 (1996) (discussing effect of different *mens rea* requirements and intended effect from using different *mens rea*).

142. *See id.* at 10 (indicating Congress’s desire to punish hackers who unintentionally cause damage to computer systems).

ris,¹⁴³ the court “accepted the government’s view that 1986 amendments to the [Computer Fraud and Abuse Act] eliminated any distinction between a break-in that damages files or steals money and what Morris was found guilty of, intentional unauthorized access that prevented authorized use.”¹⁴⁴ More recently, the court in *United States v. Sablan*¹⁴⁵ ruled that the government does not have to prove intentional damage to a computer file but only intentional access without authorization.

3. *A Critical Evaluation*

The regulatory model justifying such *expansion* of the CFAA is flawed. The most immediate weakness is that the CFAA attempts to regulate “hacking,” a particular type of computer crime without recognizing the important distinction between computer technology and the individuals responsible for its application. As described above, hacking is only one expression of computing prowess for the hacker who is equally technically proficient to engage in cracking and phreaking. Thus, to the extent that hackers engage in hacking either to express their computer prowess or to exploit structural deficiencies for monetary purposes, aggressive enforcement of hacking is unlikely to reduce the overall number of incidents of information-related crime. At best, a crackdown on hacking will prompt a shift in hackers from hacking to phreaking, cracking, or other related activities.

At worst, the crackdown on hacking represented by the CFAA is likely to prove counter-productive when analyzed from the perspective of other sources of behavioral constraint in cyberspace. For example, the public perception effectuated by the recent criminalization of all forms of hacking exacerbates the tense divide between hackers, law enforcement, and the general Internet public.¹⁴⁶ Inasmuch as laws affect the development of social norms, the average Internet user can be expected to view all forms of hacking as criminal and as undermining the consumer trust necessary for electronic commerce.

The application of public law is also allocatively inefficient in this context. The vigorous regulation of hackers through governmental law enforcement externalizes the costs of enforcing such norms to individuals

143. 928 F.2d 506 (2nd Cir. 1991).

144. Harold L. Burstyn, *Computer Whiz Guilty*, 76 A.B.A. J. 20, 20 (1990).

145. 92 F.3d 865, 865 (9th Cir. 1996).

146. For an insightful critique of current law enforcement along these lines, see Catherine Therese Clarke, *From CrimINet to Cyber-Perp: Toward an Inclusive Approach to Policing the Evolving Criminal Mens Rea on the Internet*, 75 OR. L. REV. 191 (1996).

who do not participate in electronic commerce. This unnecessary cost externalization, in turn, suggests that some form of indirect regulation through the market might best regulate hackers, at least from the perspective of allocative efficiency.¹⁴⁷

Finally, the CFAA is no different from any other example of direct regulation that has proven highly ineffective in cyberspace.¹⁴⁸ For instance, jurisdiction is problematic: foreign hackers might not be within the reach of state and federal laws,¹⁴⁹ and the complexity of Internet routing creates jurisdictional conflicts among the localities, states, and countries that wish to exercise jurisdiction over transient information packets.¹⁵⁰ In short, the myriad problems that have plagued this statute from its inception justify a wholesale rejection of its approach to regulating hackers.¹⁵¹

147. See discussion *infra* Part IV.B. Moreover, the CFAA does not provide an incentive for anyone to adopt adequate anti-hacking security measures. In fact, network security remains at an shockingly low level and is virtually nonexistent in many companies despite the severity of the hacking threat. A 1996 survey revealed that 58 percent of companies do not have a written policy on how to deal with network intrusions. See Gripman, *supra* note 18, at 174 n.21. This lack of security obviously facilitates Internet hacking. According to security expert Clifford Stoll, "The security weaknesses of both systems and networks, particularly the needless vulnerability due to sloppy systems management and administration, result in a surprising success rate for unsophisticated attacks." *Id.* at 177. This is not to say, of course, that allocative inefficiency or cost externalization is in and of itself sufficient justification for cyberspace regulation. See discussion *infra* Part V.C.

148. See Lessig, *Constitution of Code*, *supra* note 17, for a detailed discussion of Lessig's theory of indirect regulation through code as the most effective means of regulation in cyberspace.

149. See *id.* at 184.

150. See Llewellyn Joseph Gibbons, *No Regulation, Government Regulation, or Self-Regulation: Social Enforcement or Social Contracting for Governance in Cyberspace*, 6 CORNELL J.L. & PUB. POL'Y 475, 489 (1997).

151. Some proposals have suggested piecemeal reforms to existing legislation. See, e.g., Clarke, *supra* note 146. Catherine Clarke has proposed a scheme for law enforcement on the Internet that employs the technical expertise of hackers to improve Internet security while promoting self-regulation of the Internet through code solutions such as PGP. Although Clarke recognizes the importance of tailoring law enforcement techniques to match more closely available demographic data on the different subsets of the hacking community, implementing the proposals set forth thereafter are difficult to envision under the current legal regime. For instance, there is no reason to believe that convicted ex-hackers will serve as effective community educators as she suggests, particularly since the social divide between hackers and the rest of the Internet community is imposed by the law itself, irrespective of how such law is enforced. Moreover, as Clarke concedes, "cultural barriers exist between young hackers ... and police officers. Law enforcement officers may be hesitant to seek out the advice of persons who could be their teenage

B. Hacking as Tort: The Internet Service Provider ("ISP") Solution

Some academics have advanced a critique of public law solutions that implies a promising regulatory scheme based in tort negligence theory.¹⁵² These critics maintain that even if jurisdictional issues are solved, "the infrastructure of cyberspace is evolving too rapidly for governments to regulate efficiently."¹⁵³ Also, anecdotal evidence suggests that the poor enforcement record of existing criminal law will deter companies from going online. Tom Peltier, the corporate information protection coordinator for Detroit Edison Power Company, predicts that "because of the risk of online crime is so great, there will be a mass exodus of corporate users of the Internet when they realize their vulnerability."¹⁵⁴ Pointing to such deficiencies in current criminal law, proponents of a negligence regime frame their regulatory model on the following: (1) how to provide an incentive for Internet participants to increase security; (2) how to deter hacking; and (3) how to provide a financial remedy to those harmed by hacker intrusions,¹⁵⁵ both as a means to achieve the basic tort end of compensating victims and as a means of promoting online participation.

1. *The Decisive Advantages of a Negligence Regime*

Tort law does provide a more efficient means of achieving such goals. The primary purposes of tort law are: (1) to deter wrongful conduct; (2) to encourage socially responsible behavior; and (3) to compensate injured parties.¹⁵⁶ Imposing tort liability on larger market actors for losses caused by hacking encourages them to adopt socially valuable security measures (i.e., those whose expected benefits outweigh their costs); imposing tort liability on (non-judgment-proof) hackers deters them from infiltrating

children. The Generation-X young men ... may also be unenthusiastic about assisting law enforcement agencies." Clarke, *supra* note 146, at 233. Clarke must ultimately reduce her claim to the proposition that "existing institutional and procedural measures may force some level of cooperation." *Id.* Re-examination of the laws creating these harmful social norms (i.e., the social divide) suggest that existing institutional and procedural measures should be jettisoned altogether.

152. See, e.g., Gripman, *supra* note 18.

153. Gibbons, *supra* note 150, at 509.

154. Gripman, *supra* note 18, at 170 n.14.

155. See *id.* at 175. Gripman suggests imposing tort liability on corporations for injuries incurred by third parties as a result of hackers' using the corporations' networks to hack into third parties' computers. As explained below, however, such an approach would raise the cost of online participation for corporations, thereby deterring many companies—particularly small ones—from going online.

156. See *id.* at 176.

companies because of the threat of monetary sanctions.¹⁵⁷ Tort liability would prompt these market actors to exercise reasonable care in providing network security.¹⁵⁸ Parties injured by a breach of this duty are compensated for their injuries and are restored "to their original condition, insofar as the law can do this."¹⁵⁹ In this way, tort law provides incentives for such market actors to erect security measures that will lower the cost of going online to potential and current Internet participants, thereby allowing the individual expected benefits accruing to corporations from online participation to exceed the expected cost.¹⁶⁰

Numerous reasons may be cited for the proposition that ISPs should be jointly liable for the tortious hacking activities of those who use their services. As explained above, hackers are generally judgment-proof,¹⁶¹ so victims of hacking intrusions are usually left without financial remedy, thereby deterring online participation due to the high expected costs stemming from such intrusions.¹⁶² Also, hackers are difficult to detect, much less identify. Third, the jurisdictional problems discussed above render hackers very unattractive defendants. Fourth, many Internet participants also may be judgment-proof due to thin capitalization, given the low

157. *See id.* Given the difficulties associated with identifying the perpetrators of tortious hacking, the primary goal of the model of tort law proposed here is not the deterrence of socially undesirable activity (i.e., hacking), which tort law is traditionally concerned with, but rather the growth of the Internet as facilitated by greater network security.

158. *See id.*

159. *See id.* at 176 (quoting John W. Wade, et al., PROSSER, WADE AND SCHWARTZ'S TORTS 1 (9th ed. 1994)).

160. Corporations take only individual, not social, costs and benefits into account when they make business decisions. However, online participation has strong positive externalities due to such phenomena as network effects that augment the utility of other users. Thus, the social benefit of an individual corporation's online participation exceeds its individual benefit, and should therefore be encouraged. Hacking imposes a cost on online companies; compensation via tort liability reduces this cost, thereby raising the expected net benefit (benefit less cost) of going online. Thus, the tort system can raise online participation to the socially-optimal level by transferring a portion of the expected cost of going online—i.e., costs imposed by hackers—from corporations to ISPs.

161. *See* Victoria A. Cundiff, *Trade Secrets and the Internet: A Practical Perspective*, COMPUTER LAW., Aug. 1997 at 6, 14 ("Internet tortfeasors and infringers are likely to include a high percentage of students and others who may not have the resources to satisfy large judgments.").

162. ISPs may also be judgment proof in some instances. This problem could be solved by requiring ISPs to maintain a minimum level of assets.

barriers to entry on the Internet,¹⁶³ so they may be a weak source of compensation for hacking losses if they are the means by which hackers intrude into third parties' computers. Fifth, ISPs are the least cost avoiders, i.e., it is more efficient for ISPs to secure their entire systems than for each online corporation to do so, and ISPs could spread the cost of security among their subscribers. Sixth, absent liability for abuse by their subscribers, ISPs have a financial incentive to tolerate such abuse, in order that they may attract and maintain such subscribers as customers. Seventh, if corporations are forced via tort liability to provide their own security protections, the private expected cost incurred by many individual corporations on the Internet—from invoking security measures as well as from non-recoverable hacking losses—would exceed the expected gain derived from being online, thereby deterring online participation—a socially sub-optimal outcome.¹⁶⁴ This analysis is particularly applicable to small companies for whom security costs would make up a large proportion of total costs; while large corporations might be able to afford to hire security experts to constantly update their computer systems, many small companies would not. These small companies would therefore either choose not to go online (because the expected costs would outweigh the expected benefits), or would go online (e.g., if they were judgment-proof) without adequate security and threaten the security of third parties' networks with whom such companies are linked. If, however, ISPs are held liable, such costs of online activity will be transferred to the ISP, thereby increasing the private net gain of going online. Only under this regime would there exist an adequate incentive to adopt security measures without sacrificing online participation and growth.

The negligence rule¹⁶⁵ thus provides an allocation of incentives, deterrence, and remedies as it fulfills three primary objectives: (1) the provision

163. See Ian C. Ballon, *The Law of the Internet: Developing a Framework for Making New Law*, 482 PLI/PAT 9, 20-21 (1997).

164. One might argue that a corporation's knowing placement of confidential information in a database accessible on the Internet constitutes an effective assumption of risk that would vitiate third party tort liability. However, unlike other risky activities (e.g., skiing), online activities have positive externalities and should be encouraged, given the network effects of online participation and the efficiency of electronic commerce. Tort liability imposed on ISPs largely removes such risk from corporations' net benefit calculus and therefore increases their expected net benefit from online participation, thereby increasing total expected online participation.

165. Strict liability is another regime that could be possibly erected to deal with the hacking problem. Applying strict liability to ISPs for all damages incurred as a result of hacking has its advantages, given that (1) ISPs are the party in the best position to detect

of an incentive for ISPs to augment their security levels; (2) the deterrence of (non-judgment-proof) hackers from illegally breaking into computer networks because even unintentional harm may make them liable; and (3) the provision to injured corporations or persons a financial remedy for their injuries.¹⁶⁶

2. A Critical Evaluation

Although a tort-based model of regulation is more efficient than the status quo, an examination of the process by which an ISP liability system would seek to regulate hackers suggests several compelling bases for nevertheless rejecting such a scheme. Cost-benefit analysis in such a legal regime would arguably create incentives for Internet participants to deter hacking at the level of Internet technology or architecture. In fact, proponents of a tort-based model of reform criticize inadequate network security¹⁶⁷ and suggest that a minimum security standard for determining the duty of reasonable care would include the adoption of architectural or code solutions¹⁶⁸ such as encryption technology and Internet Protocol next generation ("IPng"), widely touted as "the future version of IP used on the Internet ... [providing] support for authentication, data integrity, and confidentiality."¹⁶⁹ In short, the ISP liability model advocates a shift in the form of behavioral constraint from "direct" regulation of hacking activity to "indirect" regulation through Internet code or architecture.¹⁷⁰

and eliminate defects in security, (2) ISPs are best able to absorb and spread the risk or cost of injuries through insurance or price increases, and (3) the strict liability rule avoids costly and burdensome requirements of proof. However, the problem with such an approach is that it limits online corporations' incentives to establish security systems of their own that exceed the security levels imposed on ISPs by a due care standard, since corporations would be compensated for all losses regardless of whether the ISP maintained the level of due care or not. Under the negligence rule, this would not be a problem. If a corporation felt that the level of due care was too low for its purposes (say, because it had unusually highly sensitive and valuable information exposed), it would have an incentive to erect higher security levels than those required under due care, since the corporation would not be compensated for losses if the ISP maintained the level of security mandated under the due care standard.

166. See Gripman, *supra* note 18, at 179.

167. See *id.* at 171-77.

168. See *id.* at 184-91.

169. William A. Hodkowski, *The Future of Internet Security: How New Technologies Will Shape the Internet and Affect the Law*, 13 SANTA CLARA COMPUTER & HIGH TECH. L.J. 217, 220 (1997).

170. See generally Lawrence Lessig, *Constitution of Code*, *supra* note 17.

However, such reliance on technology-based solutions is problematic in two important respects. First, although most “live life subject to the constraints of code ... however (and by whomever) these constraints have been set,”¹⁷¹ hackers do not. To be sure, technological constraints on hackers may have some impact on network security, but the regulatory implication drawn from the history of Internet technology is clear: any such impact will be temporary and inadequate for the purposes of securing electronic commerce.¹⁷² In this instance, the government cannot accomplish indirectly that which technology precludes it from doing directly.

Second, and perhaps more importantly, proponents of the ISP liability scheme have failed to consider the broader policy implications of regulation that has the effect of altering cyberspace code. As Lessig has argued, the architecture of the Internet has the potential to enable and disable behavior with a level of “efficiency” and “compliance” impossible to achieve in real space.¹⁷³ The ability to raise such powerful *ex ante* mechanisms of control, in turn, raises public policy concerns even though imposing ISP liability would permit the market to determine changes in code. The ISP model does empower consumers to approve or reject changes in the code, but any such change would be

the result of a collection of choices made at an individual level, [with] no collective choice made at a collective level. It is the product of the market. But individual choice might aggregate in a way that individuals collectively do not want. Individual choices are made within a particular architecture; but they may yield an architecture different from what the collective might want.¹⁷⁴

Particularly since no code-based solution is likely to regulate hackers effectively, the danger posed by making uninformed changes in code premised on the regulation of hackers may indeed be greater for the vast majority of Internet users than any of the current dangers posed by hackers.

171. *Id.* at 184.

172. See discussion *supra* Part II. The proposition that hackers will evade architectural constraints and therefore pose a threat to electronic commerce is distinct from the claim that hackers point to the general inefficacy of code-based solutions in cyberspace, an argument which is not made here.

173. See Lessig, *Constitution in Cyberspace*, *supra* note 17, at 869.

174. Lawrence Lessig, *The Zones of Cyberspace*, 48 STAN. L. REV. 1403, 1411 (1996) [hereinafter *Zones*].

V. A HEURISTIC MODEL FOR REFORM

This critical survey of federal legislation and academic scholarship has attempted not only to expose the critical problems with current models for regulating hackers, but also to deconstruct the current discourse so as to expose the underlying sources of widespread regulatory failure. Such an approach has produced insights that inform the development of a regulatory framework perhaps better equipped to address the threat posed by hackers to electronic commerce. This framework, in turn, should provide the basis for a fruitful discussion of the broader respects in which cyberspace has transformed social dynamics.

A. Consider All Relevant Modalities of Regulation

Although several distinct forms of regulation have emerged in the context of cyberspace, neither current legislation nor the academic literature on hacking has produced a regulatory model that fully comprehends the effects of multiple feedback loops between the several sources of constraint through which social behavior may be regulated, both in real space and in cyberspace. Professor Lessig has labeled these sources of behavioral constraint “the modalities of regulation” and identified the primary modalities as “code” (or architecture/geography), “law,” “social norms” and the market.¹⁷⁵ According to Lessig:

[L]aw, norms and code regulate cyberspace just as law, norms and nature (or what I call “real space code”) regulate real space. But there is an important difference between these two regimes. In real space, constraints are changed by changing law; in cyberspace, constraints will be changed by changing code. This will follow because of two features of these two different worlds: First: In real space, it is law that is plastic; in cyberspace, it is code that is plastic. And second: In real space, it is relatively hard to escape the constraints of law; in cyberspace, it is much easier. The effect of both differences will be to shift the locus of regulatory change from law to code. In real space, law is at cen-

175. See Lawrence Lessig, *Constitution of Code*, *supra* note 17. Although Lessig makes explicit reference only to code, law, and social norms, he does not claim “that there are no other constraints. Psychology or the market, for example, are constraints which are related to these three primary constraints in complex ways.” *Id.* at 181 n.1. Explicit mention of market forces above is consistent with Lessig’s inclusion of the market as a primary constraint in his more recent lectures in his course *The High Tech Entrepreneur*.

ter stage, and code is an afterthought. In cyberspace, the game is code. Law is a side-show.¹⁷⁶

Viewed through the lens of Lessig's framework, recent criticisms of the CFAA and even more recent proposals for imposing tort liability on market actors are part of a broader paradigm shift in cyberspace regulation. The CFAA is an attempt to regulate hackers directly through law, and recent scholarship advancing the "indirect regulation" of code through market manipulation in lieu of the CFAA confirms Lessig's theory that "the locus of regulatory change" is changing. Indeed, governmental adaptation of such proposals may not be far behind.¹⁷⁷ And yet, any such "indirect regulation" would be misguided insofar as it attempted to regulate hackers. In Lessig's terms, "Hackers define for themselves a certain anarchy, by devoting themselves to finding the holes in the existing code."¹⁷⁸

It does not necessarily follow, however, that hackers are impossible to regulate. Lessig's articulation of the primacy of code-altering regulation in cyberspace is a predictive model to facilitate critical analysis of a general trend, not a prescriptive model for how to regulate cyberspace effectively in every instance. In fact, the very process of modality-interplay by which code becomes "the game" suggests that code does not have to be the sole conduit of cyberspace regulation. As Lessig points out, "Architectures don't come in natural kinds."¹⁷⁹ Rather, Internet architectures reflect choices, ones that have been encoded with the values informing those choices—and vice versa. In this respect, code operates not only as an *ex*

176. *Id.* at 183-84 (footnotes omitted).

177. *See id.* at 184.

[G]overnment will shift to a different regulatory technique. Rather than regulating behavior directly, government will regulate indirectly. Rather than making rules that apply to constrain individuals directly, government will make rules that require a change in code, so that code regulates differently. Code will become the government's tool. Law will regulate code, so that code constrains as government wants.

Id.

178. Lessig, *Zones*, *supra* note 174, at 408 n.18. Lessig's contention that indirect regulation through code is the most effective regulator in cyberspace in no way competes with the contention that such code is a poor means for regulating hackers. Lessig's fear is that cyberspace code will develop in undesirable ways despite the existence of hackers, not as a consequence of eliminating hackers. ("I don't think one need believe hacking impossible to believe it will become less and less significant. People escaped from concentration camps, but that hardly undermines the significance of the evil in concentration camps."). *Id.*

179. Lessig, *Constitution of Code*, *supra* note 17, at 1411.

ante constraint on socially undesirable behavior, but also by way of its relation to social meaning—as a code of ethics or social norms.¹⁸⁰

This process in turn exposes a powerful framework for applying the interplay between modalities for prescriptive purposes. In this respect, the strength of Lessig's model is the ease with which modalities may be viewed for their effects on each other. Such interplay points to alternative modalities of regulation for hackers, who cannot be regulated as others in cyberspace are. Admittedly, direct regulation through law (e.g., the CFAA) has been as unsuccessful as code-altering solutions. Nevertheless, examining the effect of law in general, particularly through its interplay with other modalities, may yield answers, whereas exhaustive exploration of the nature of Internet technology has yielded none.

B. Analyze the Political Consequences of Inducing Changes In Code

Equally, if not more important, than which modalities regulate social behavior is the issue of who controls those modalities, whether in real space or cyberspace. In cyberspace, code takes on many of the characteristics that make law effective in real space. With respect to its power to alter social behavior, then, the architecture of the Internet (i.e., the regulation thereof) is more properly the analog of real space law than real space geography or architecture. Yet real space law and cyberspace code also differ in ways that suggest the need for careful scrutiny of code-altering regulation.

The most striking difference between real space law and cyberspace code is that law regulates “through the threat of ex post sanction, while code, in constructing a social world, regulates immediately.”¹⁸¹ However, the most visible instances of code's immense regulatory power take the form of “zoning” technologies,¹⁸² or commercial alterations of code designed to create “a perfect technology of choice”¹⁸³ for Internet users (e.g., by making each inhabitant of cyberspace “a market of one”). That government might indirectly regulate the market to induce such changes for its regulatory purposes is less clear to individual users. Yet government can easily transform commercially designed code into “a perfect technology of justice,” one that allows policymakers to select a social end, and

180. Initially, these ethics reflected the values of Internet architects. This is certainly not the case today. See discussion *supra* Part IV.B.

181. Lessig, *Constitution of Code*, *supra* note 17, at 184.

182. See Lessig, *Constitution in Cyberspace*, *supra* note 173, at 901.

183. Lessig, *Zones*, *supra* note 174, at 1410.

then assure compliance by individuals to that end.¹⁸⁴ It is in this profound sense that code is of great political consequence in and of itself, regardless of whether the government or the market is shaping its development:

“[S]tructures of [code-altering] regulation entail important value choices. Whether information will be kept private, whether encrypted speech is allowed, whether anonymity is permissible, whether access is open and free—these are policy choices made by default by a structure of code that has developed—unaware at times, and, generally, uncritically of the politics that code entails.”¹⁸⁵

Thus, code in cyberspace possesses regulatory power far beyond the reach of law in real space, and yet alterations of code currently do not undergo any scrutiny resembling the democratic process by which laws are legitimated in real space.¹⁸⁶ Of course, commercial code developers and scholars alike continue to propose changes, such as imposing tort liability on market actors as an “efficient” form of inducing changes in the code to advance the goal of Internet security. Such proponents fail to see the broader issue of choice:

We could imagine allowing efficiency to rule this new space, by allowing liberties protected by imperfections to fall away; or we could imagine recreating spheres of liberty to replace those created by imperfections in technology. These are our democratic choices, and real choices they are.¹⁸⁷

That market actors fail to address these choices is acceptable, perhaps inevitable; that policymakers and academics do so is irresponsible.

VI. A PROPOSAL FOR OPTIMAL REGULATION

The preceding heuristic model suggests that state and federal governments should immediately decriminalize all forms of non-malicious hacking. Non-malicious hacking should be defined as obtaining unauthorized access to a protected computer without causing intentional or reckless damage. Successful incidents of unauthorized access should be presumed by law to be non-malicious if the actor makes a good-faith effort

184. *Id.* at 1408.

185. Lessig, *Constitution of Code*, *supra* note 17, at 184.

186. *See* Lessig, *Zones*, *supra* note 174, at 1410.

187. Lessig, *Constitution in Cyberspace*, *supra* note 173, at 909.

to report the incident to the proprietor of the accessed system immediately upon obtaining access.

All existing state and federal statutes governing computer-related activities, including the Computer Fraud and Abuse Act ("FCAA"), 18 U.S.C. § 1030 (1994), should be amended to reflect this change in policy. In particular, Section 1030(a)(5)¹⁸⁸ of the FCAA should be modified to repeal the third provision criminalizing acts causing negligent damage by outside actors.

A. Advantage One: Promotes Self-Regulation Through Market Forces

The proposed change in legislation enables market actors to draw upon the resources of non-malicious hackers to increase the security level of the Internet and otherwise to mitigate the economic threat posed by malicious hackers. Currently, all hackers are treated by the law and viewed by the public as dangerous criminals who must be stopped at all costs. This over-generalization discourages companies and law enforcement agents from enlisting the help of hackers in identifying latent security flaws and collecting information on acts of malicious hacking. Moreover, the sweeping criminalization of all hacking activities has bred within the hacking community a strong distrust and resentment of computer security professionals and government agents.¹⁸⁹

A clear legal distinction between malicious and non-malicious hacking will revive the positive and self-regulating norms (i.e. the "hacker ethic") within the hacking community and promote market-based initiatives aimed at enlisting the help of non-malicious hackers.¹⁹⁰ Existing literature indicates that many within the hacking community would be willing to cooperate with companies and government agencies if monetary rewards and public recognition were offered for their skills and knowledge.¹⁹¹ Such market-based initiatives may include the following:¹⁹²

188. See discussion *supra* Part IV.A.

189. Hackers felt that system managers treat them like enemies and criminals, rather than as potential helpers in their task of making their systems secure. See Dorothy E. Denning, *Concerning Hackers Who Break into Computer Systems* (visited Apr. 24, 1999) <<http://www.cpsr.org/cpsr/privacy/crime/denning.hackers.html>>.

190. "Frank Drake," an editor of the now defunct cyberpunk W.O.R.M., suggested in 1990 that making a legal distinction between malicious and non-malicious hacking would lead to a "kinder, gentler" relationship between hackers and computer security people. See *id.* at 16.

191. According to Dorothy Denning in her 1990 survey, several hackers said that they would like to be able to pursue their activities legally and for income: "Hackers say

- 1) Companies can offer monetary rewards and public recognition for hackers who voluntarily report their successful break-ins and give suggestions for correcting latent security flaws.¹⁹³
- 2) Companies and government agencies can offer monetary rewards for hackers who provide useful information about acts of malicious hacking.¹⁹⁴
- 3) Companies can hire hackers as security consultants or members of "tiger teams."¹⁹⁵

By tapping into the expertise and knowledge base of hackers, who are often in the best position to identify security holes, companies can receive invaluable assistance in detecting and correcting latent security flaws before they are exploited for malicious purposes. In addition, the public will have a more positive view of hackers in general, increasing consumer trust

they want to help system managers make their systems more secure. They would like managers to recognize and use their knowledge about design flaws and the outsider threat problem." Also, the hackers felt that it would help if system managers and the operators of phone companies and switches could cooperate in tracing a hacker without bringing in law enforcement authorities. *See id.* at 15.

192. As the following footnotes will illustrate, some companies are turning to market-based initiatives already. With the decriminalization of non-malicious hackers, more and more companies will feel comfortable with trusting hackers and relying on them for their expertise.

193. For example, consider Crypto-Logic. This company has developed a new type of encryption software for sending secure e-mail messages. It is currently staging a contest in which it challenges hackers to decode an encrypted message sitting on its Web site. *See Ultimate Privacy* (visited Feb. 8, 1999) <<http://www.ultimateprivacy.com>>.

194. For instance, in the famous case of Rome Laboratory Attacks, the Government was able to identify one of the hackers through an intelligent network of informants after failed attempts to trace back the origin of attack using phone taps and packet tracing tools. *See Christy, supra* note 79, at 59-60.

195. Although the information security community is in principle reluctant to hire hackers to work for them, some will admit to hiring, or at least consulting with, ex-hackers. Among them are the National Computer Crime Information Center, part of the Federal Bureau of Investigation, and the operator of the system that is hacking's Holy Grail: the National Security Agency ("NSA"). A highly regarded Information Services security consultant confirmed that both institutions, along with several major defense contractors, have occasionally used hackers at least as informants in the past.

In another instance, Price Waterhouse's elite group of computer experts—the Tiger Team—spends its waking hours breaking into their client's security systems. The team, part of the firm's Enterprise Security Solutions Practice, simulates "enemy" break-ins to help clients defend themselves against computer hackers.

in the safety of conducting commercial transactions online.¹⁹⁶ Finally, government agencies can make better use of law enforcement resources by focusing on deterring and prosecuting malicious hackers.

B. Advantage Two: Facilitates Democratization of Architectural Developments

In addition to mitigating the economic threat posed by malicious hackers, the proposal also facilitates an informed discussion of the political nature of code or the “governance” issue implicit in code-altering regulation. As mentioned above, changes in code are currently implemented without any formal institution or process for review, recommendation, or legitimization. Governmental and commercial code developers, whose interests are often aligned in this respect, thus possess potentially unchecked discretion in their development of Internet architecture. However, the decriminalization of non-malicious hacking presents an opportunity to place a check on corporate and governmental interests and initiate a scheme for the democratization of Internet code development.

Just as the free flow of market forces in a proposed regime of decriminalization would forge the necessary “trust” between consumers and retailers to promote the growth of electronic commerce, decriminalization of non-malicious hacking also bridges the cultural gap between hackers and the vast majority of Internet users. Without the force of law to create a widespread societal norm against the activities of non-malicious hacking, both hackers and ordinary Internet users are likely to find their interests aligned. Should a group of hackers organize for the purposes of promoting open discussion of the policy implications of adopting various changes in code, there is no reason to believe consumers would hesitate to pay attention, particularly where the advice of hackers was contrary to that offered by corporate and governmental forces. In this respect, hackers could become at the very least a loosely organized coalition of consumer advocates who could provide a forum, however informal, for the discussion and implementation of code at a collective level. In fact, many existing hacker organizations could fulfill such a function; nor does one organization have to operate in such a capacity to the exclusion of others.¹⁹⁷ Consumers may ultimately make the same “choice” they would have made without the quasi-institutional role played by hackers. But under the presence of their

196. See discussion *supra* Part I.

197. In fact, many hackers are members of consumer advocate and civil liberties organizations such as Electronic Frontier Foundation (“EFF”), the League for Programming Freedom (“LPF”), and SotMesc.

watchful eyes, the process of implementing architectural changes in cyberspace will more likely reflect the democratic principles that govern this nation in real space.

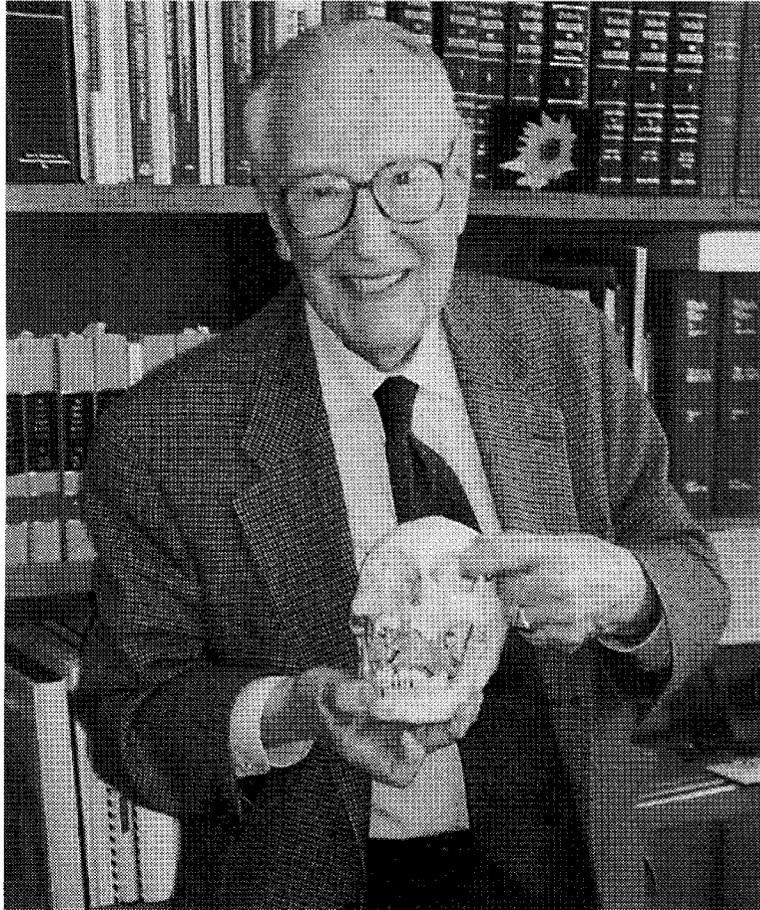


Photo Courtesy of Janice Mueller

JUDGE GILES S. RICH
1904–1999

IN MEMORIAM JUDGE GILES S. RICH

Judge Giles S. Rich, considered by many scholars to be the father of modern patent law, passed away on June 9, 1999. Judge Rich was a historical and active force in shaping this country's intellectual property system, both in the legislature and on the bench. A member of the Drafting Committee of the Coordinating Committee of the National Council of Patent Law Associations, he co-authored the 1952 Patent Act, which remains the basis of the current patent law.

Judge Rich was also active in reform from the bench. During his lifetime, Judge Rich had the distinction of being the oldest active federal judge in the history of the United States, serving on the United States Court of Customs and Patent Appeals from 1956 to 1982 and then on the Federal Circuit from its inception in 1982 until his death this past year. He was the author of many major panel and en banc decisions, including *In re Alappat* and *State Street Bank & Trust Co. v. Signature Financial Group, Inc.*, two decisions primarily responsible for opening up the patent system to software and Internet business methods.

The Berkeley Technology Law Journal is pleased to present the following memorials: one from Judge Paul R. Michel, a long-time Federal Circuit colleague of Judge Rich, and two from former law clerks of Judge Rich, Neil A. Smith and Janice M. Mueller. As patent law moves forward into the 21st century, the legacy of Judge Rich will move with us, guided by his efforts during over fifty-plus years of legislative and judicial scrutiny.

