

Production: Produced by members of the *Berkeley Technology Law Journal*.
All editing and layout done using Microsoft Word.

Printer: Joe Christensen, Inc., Lincoln, Nebraska.
Printed in the U.S.A.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Library Materials, ANSI Z39.48—1984.

Copyright © 2012 Regents of the University of California.
All Rights Reserved.

Berkeley Technology Law Journal
University of California
School of Law
3 Boalt Hall
Berkeley, California 94720-7200
btlj@law.berkeley.edu
<http://www.btlj.org>



BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 27

NUMBER 3

SYMPOSIUM 2012

TABLE OF CONTENTS

SYMPOSIUM

ORPHAN WORKS & MASS DIGITIZATION: OBSTACLES & OPPORTUNITIES

KEYNOTE ADDRESS: ORPHAN WORKS & MASS DIGITIZATION: OBSTACLES & OPPORTUNITIES	1251
<i>Maria A. Pallante</i>	
PRIVATE DIGITAL LIBRARIES AND ORPHAN WORKS.....	1259
<i>Randal C. Picker</i>	
THE ORPHANS, THE MARKET, AND THE COPYRIGHT DOGMA: A MODEST SOLUTION FOR A GRAND PROBLEM.....	1285
<i>Ariel Katz</i>	
THE ORPHAN WORKS CHIMERA AND HOW TO DEFEAT IT: A VIEW FROM ACROSS THE ATLANTIC	1347
<i>Stef van Gompel</i>	
HOW FAIR USE CAN HELP SOLVE THE ORPHAN WORKS PROBLEM.....	1379
<i>Jennifer M. Urban</i>	
ABANDONING THE ORPHANS: AN OPEN ACCESS APPROACH TO HOSTAGE WORKS.....	1431
<i>Lydia Pallas Loren</i>	
ATOMISM AND AUTOMATION.....	1471
<i>Molly Shaffer Van Houweling</i>	
ORPHAN WORKS AS GRIST FOR THE DATA MILL	1503
<i>Matthew Sag</i>	

SUBSCRIBER INFORMATION

The *Berkeley Technology Law Journal* (ISSN1086-3818), a continuation of the *High Technology Law Journal* effective Volume 11, is edited by the students of the University of California, Berkeley, School of Law (Boalt Hall) and is published in print three times each year (March, September, December), with a fourth issue published online only (July), by the Regents of the University of California, Berkeley. Periodicals Postage Rate Paid at Berkeley, CA 94704-9998, and at additional mailing offices. POSTMASTER: Send address changes to Journal Publications, University of California, Berkeley Law—Library, LL123 Boalt Hall—South Addition, Berkeley, CA 94720-7210.

Correspondence. Address all correspondence regarding subscriptions, address changes, claims for non-receipt, single copies, advertising, and permission to reprint to Journal Publications, University of California, Berkeley Law—Library, LL123 Boalt Hall—South Addition, Berkeley, CA 94705-7210; (510) 643-6600; JournalPublications@law.berkeley.edu. *Authors:* see section titled Information for Authors.

Subscriptions. Annual subscriptions are \$65.00 for individuals and \$85.00 for organizations. Single issues are \$30.00. Please allow two months for receipt of the first issue. Payment may be made by check, international money order, or credit card (MasterCard/Visa). Domestic claims for non-receipt of issues should be made within 90 days of the month of publication; overseas claims should be made within 180 days. Thereafter, the regular back issue rate (\$30.00) will be charged for replacement. Overseas delivery is not guaranteed.

Form. The text and citations in the *Journal* conform generally to the THE CHICAGO MANUAL OF STYLE (16th ed. 2010) and to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 19th ed. 2010). Please cite this issue of the *Berkeley Technology Law Journal* as 27 BERKELEY TECH. L.J. ____ (2012).

BTLJ ONLINE

The full text and abstracts of many previously published *Berkeley Technology Law Journal* articles can be found at <http://www.btlj.org>. Our site also contains a cumulative index, general information about the *Journal*, and the Bolt, a collection of short comments and updates about new developments in law and technology written by members of BTLJ.

INFORMATION FOR AUTHORS

The Editorial Board of the *Berkeley Technology Law Journal* invites the submission of unsolicited manuscripts. Submissions may include previously unpublished articles, essays, book reviews, case notes, or comments concerning any aspect of the relationship between technology and the law. If any portion of a manuscript has been previously published, the author should so indicate.

Format. Submissions are accepted in electronic format through the ExpressO online submission system. Authors should include a curriculum vitae and resume when submitting articles, including his or her full name, credentials, degrees earned, academic or professional affiliations, and citations to all previously published legal articles. The ExpressO submission website can be found at <http://law.bepress.com/expresso>.

Citations. All citations should conform to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 19th ed. 2010).

Copyrighted Material. If a manuscript contains any copyrighted table, chart, graph, illustration, photograph, or more than eight lines of text, the author must obtain written permission from the copyright holder for use of the material.

DONORS

The *Berkeley Technology Law Journal* and the Berkeley Center for Law & Technology acknowledge the following generous donors to Berkeley Law's Law and Technology Program:

Partners

COOLEY LLP	ORRICK, HERRINGTON & SUTCLIFFE LLP
FENWICK & WEST LLP	

Benefactors

COVINGTON & BURLING LLP	SKADDEN, ARPS, SLATE, MEAGHER & FLOM LLP & AFFILIATES
FISH & RICHARDSON P.C.	WEIL, GOTSHAL & MANGES LLP
KASOWITZ BENSON TORRES & FRIEDMAN LLP	WHITE & CASE LLP
KIRKLAND & ELLIS LLP	WILMER CUTLER PICKERING HALE AND DORR LLP
LATHAM & WATKINS LLP	WILSON SONSINI GOODRICH & ROSATI
MCDERMOTT WILL & EMERY	WINSTON & STRAWN LLP
MORRISON & FOERSTER LLP	

Members

ALSTON + BIRD LLP	KILPATRICK TOWNSEND & STOCKTON LLP
BAKER BOTTS LLP	KNOBBE MARTENS OLSON & BEAR LLP
BINGHAM MCCUTCHEN LLP	MORGAN, LEWIS & BOCKIUS LLP
DEWEY & LEBOEUF LLP	MUNGER, TOLLES & OLSON LLP
DURIE TANGRI LLP	ROPES & GRAY LLP
FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP	SCHWEGMAN LUNDBERG WOESSNER
GUNDERSON DETTMER STOUGH VILLENEUVE FRANKLIN & HACHIGIAN, LLP	SIDLEY AUSTIN LLP
HAYNES AND BOONE, LLP	VAN PELT, YI & JAMES LLP
HICKMAN PALERMO TRUONG BECKER, LLP	WEAVER AUSTIN VILLENEUVE & SAMPSON, LLP
KEKER & VAN NEST LLP	

Patrons

BAKER & MCKENZIE

BTLJ ADVISORY BOARD

ROBERT BARR

*Executive Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

ROBERT C. BERRING, JR.

Walter Perry Johnson Professor of Law
U.C. Berkeley School of Law
Berkeley, California

JESSE H. CHOPER

Earl Warren Professor of Public Law
U.C. Berkeley School of Law
Berkeley, California

PETER S. MENELL

*Professor of Law and Faculty Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

ROBERT P. MERGES

*Wilson Sonsini Goodrich & Rosati Professor
of Law and Technology and Faculty Director of
the Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

REGIS MCKENNA

Chairman and CEO
Regis McKenna, Inc.
Palo Alto, California

DEIRDRE K. MULLIGAN

*Clinical Professor and Faculty Director of the
Berkeley Center for Law and Technology*
U.C. Berkeley School of Information
Berkeley, California

JAMES POOLEY

*Deputy Director General of the
World Intellectual Property Organization*
Washington, D.C.

MATTHEW D. POWERS

Tensegrity Law Group, LLP
Redwood Shores, California

PAMELA SAMUELSON

*Professor of Law & Information
and Faculty Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

LIONEL S. SOBEL

*Professor of Law and Director of the
International Entertainment & Media Law
Summer Program in London, England*
Southwestern University School of Law
Los Angeles, California

LARRY W. SONSINI

Wilson Sonsini Goodrich & Rosati
Palo Alto, California

MICHAEL STERN

Cooley LLP
Palo Alto, California

MICHAEL TRAYNOR

Cobalt LLP
Berkeley, California

THOMAS F. VILLENEUVE

Gunderson Dettmer Stough Villeneuve
Franklin & Hachigian LLP
Redwood City, California

BOARD OF EDITORS

2012–2013

Executive Committee

Editor-in-Chief
DAVID ROSEN

Managing Editor
LAUREN ESCHER

Senior Articles Editors
YVONNE LEE
HANNAH MINKEVITCH

Senior Executive Editor
JOE SEXTON

Senior Annual Review Editors
WINNIE HUNG
JANE LEVICH

Senior Scholarship Editor
COURTNEY BOWMAN

Editorial Board

Submissions Editors
JULIA KOLIBACHUK
PENNI TAKADE

Production Editor
JARAD BROWN

Bluebook Editors
GABRIEL MILITELLO
YUE WANG
ZACH WOOD

Annual Review Editors
ANGEL DIAZ
LALI MADDURI

Notes & Comments Editors
ELISE EDLIN
BRITT HARWOOD

Symposium Editors
KAREN KOPEL
KILEY WONG

Publishing Editor
DINA ROUMIANTSEVA

Web Content Editor
CHRIS CIVIL

Web & Technology Editors
EDUARD MELESHINSKY
ANDREA YANKOVSKY

External Relations Editor
GAVIN LIU

Member Relations Editor
CASEY HULTIN

Articles Editors
BRADY BLASCO
JULIE BYREN
EMILY CHEN
ROSS COHEN
CHRISTINA FARMER

Articles Editors
GRANT GARBER
ROBIN KUNTZ
NICHOLAS LAMPROS
MARK LANGER

Articles Editors
MARIENNA MURCH
JUSTIN ORR
SARAH ORRICK
LAUREN SMITH
ETHAN WEINER

BERKELEY CENTER FOR LAW & TECHNOLOGY 2012–2013

Executive Director

ROBERT BARR

Faculty Directors

PETER MENELL
ROBERT MERGES

DEIRDRE MULLIGAN
PAMELA SAMUELSON
PAUL SCHWARTZ

SUZANNE SCOTCHMER
MOLLY VAN HOUWELING

Associate Director

LOUISE LEE

Assistant Director

JULIA TIER

Affiliated Faculty and Scholars

AARON EDLIN
JOSEPH FARRELL
RICHARD GILBERT
BRONWYN HALL
THOMAS JORDE
MICHAEL KATZ
DAVID MOWERY

DAVID NIMMER
DANIEL RUBINFELD
ANNALEE SAXENIAN
JASON SCHULTZ
HOWARD SHELANSKI
CARL SHAPIRO

MARJORIE SHULTZ
LON SOBEL
TALHA SYED
DAVID TEECE
JENNIFER M. URBAN
HAL R. VARIAN
DAVID WINICKOFF

ORPHAN WORKS & MASS DIGITIZATION: OBSTACLES & OPPORTUNITIES

Maria A. Pallante[†]

I am honored to deliver the David Nelson lecture and I want to thank the Berkeley Center for Law and Technology for the invitation and for this outstanding symposium. I would like to revisit some of the early points of tension in the orphan works debate and the shifting assumptions about the problem, and then provide some insight into the goals of the Copyright Office during the 113th Congress.

I am aware that many people, including many in this room have studied, debated, litigated, and negotiated issues relating to orphan works, and I have worked closely with some of you in my own work. There are, of course, some differences of opinion that we will need to reconcile if we are to move from a holding pattern to progress on orphan works. That said, I do think we have a significant amount of agreement not only in this room, but also in the broader copyright community, to a degree that simply was not the case before.

For example:

- 1) We seem to have general agreement that in the case of a true orphan work, where there is no copyright owner and therefore no beneficiary of the copyright term, it does not further the objectives of the copyright system to deny use of the work, sometimes for decades. In other words, it is not good policy to protect a copyright when there is no evidence of a copyright owner. Former Chairman of the Senate Subcommittee on Intellectual Property of the Senate Committee on the Judiciary Orrin Hatch put it this way: “In other areas of law, the rights of an owner are limited by things like adverse possession, salvage rights, found treasure doctrines, and abandonment theories.

[†] Register of Copyrights and Director, U.S. Copyright Office. Ms. Pallante delivered the David Nelson Memorial Keynote Address on April 12, 2012, at the Orphan Works & Mass Digitization: Obstacles & Opportunities Symposium, sponsored by the Berkeley Center for Law & Technology and the Berkeley Technology Law Journal.

Many believe that a similar concept should be applied to copyright law.”¹

- 2) We seem to have general agreement that, even where a good faith, diligent search is undertaken (with all of the expertise and resources required), the act of finding one or more copyright owners does not always lead to resolution. On the contrary, for many users, finding one or more copyright owners is not the end, but the beginning of effort and expense, and therefore difficult to justify as a strategy for certain projects, provided we can define those projects responsibly.
- 3) Finally, I believe we agree that neither of these points, orphans or market gridlock, are good for the copyright system. It is not good for the users and authors who would otherwise engage in transactions, but perhaps more importantly, it does not engender faith in the operation of the law or respect for the goals of the law. This fact was reflected in the Copyright Principles Project, chaired by Pam Samuelson, as was the realization that to do nothing is also not an option. The group wrote that there is “no one ‘silver bullet’ that can relieve all of the stress, maintain or renew public confidence in copyright, and bring calm to copyright industries disrupted by new technologies,” but “a better copyright law is possible.”²

With these points in mind, I think that confirming, and in some cases resetting, the copyright relationship as between copyright owners and libraries and archives is of special importance to the public and therefore of primary importance to my Office. This will require hard work, a commitment from all sides to come to the table, and a willingness to compromise. In the end, we should enact orphan works legislation, update library exceptions, facilitate new licensing models, reconcile competition and pricing issues with the copyright marketplace, and protect the doctrine of fair use. I believe legislation is an important component of the solution because I think that Congress has a responsibility to update the copyright law.

Before I go any further, I want to note the importance of the academy in the orphan works debate. Schools like Berkeley have many advantages over

1. *Orphan Works: Proposals for a Legislative Solution: Hearing Before the Subcomm. on Intellectual Property of the S. Comm. on the Judiciary*, 109th Cong. 2 (2006) (statement of Sen. Orrin G. Hatch, Chairman, Subcomm. on Intellectual Property of the S. Comm. on the Judiciary).

2. Pamela Samuelson, *The Copyright Principles Project: Directions for Reform*, 25 BERKELEY TECH. L.J. 1176, 1178, 1179 (2010).

the government when it comes to problem solving, including resources, political distance, and cross-disciplinary expertise. With respect to the orphan works issue, many also have direct experience, whether through library digitization projects or academic publications. Partnering with academic institutions (whether formally or informally) is important to me as Register: it is one of ten core commitments I announced last October in the publicly released document, *Priorities and Special Projects of the Copyright Office*.³ I have formed a special cross-disciplinary committee of staff dedicated to exploring relationships with all kinds of schools around all kinds of topics, including research projects, workshops, fellowships and clerkships, and advisory committees. If this is something that may appeal to you, I hope you will let me know.

In thinking about my remarks today, I noted with interest that this is the 100th anniversary of Berkeley's School of Law and the motto is "Big Ideas, Bold Action." This seems a perfect motto for the kind of work that happens regularly at Berkeley—but I am a policy lawyer in Washington, DC.

To be clear, there is no shortage of big ideas in Washington, generally speaking. But in terms of copyright policy, because the environment is so often polarized, it is quite common for people to dismantle big ideas into smaller ones for the purpose of making them manageable. And it is frequently the case that legislative action responds to technological change, but does not anticipate it. In part this is because legislation is hard to enact, and in part it is because those on the cutting edge of technology have always been highly valued by society. As one witness told Congress at a copyright hearing nearly thirty years ago, "When you are legislating on the cutting edge of technology, the all-important thing is to stay behind the blade."⁴

Within the Copyright Office, my colleagues and I are committed to orphan works legislation and would very much like to see it as a major point of discussion in the 113th of Congress. The end of the second session will mark ten years from the date the Copyright Office began its public process, with the 2005 notice of inquiry. This may seem like a long time (particularly for those of you who are young people), and it is in fact a long time. But experienced Members of Congress will tell you that good copyright legislation takes multiple Congresses to enact. It requires debate, reflection, refinement, and compromise. But that is not the entire story. It is also true

3. U.S. COPYRIGHT OFFICE, PRIORITIES AND SPECIAL PROJECTS OF THE UNITED STATES COPYRIGHT OFFICE, OCTOBER 2011–OCTOBER 2013 (2011).

4. *Copyright and Technological Change: Hearing Before the Subcomm. on Courts, Civil Liberties, and the Admin. of Justice of the H.R. Comm. on the Judiciary*, 98th Cong. 169 (1983) (summary of Rapporteur Paul Goldstein) (quoting Haines Gaffner).

that sometimes it simply takes time for people to calm down and to get comfortable with big ideas. This may be especially true of copyright law, according to former Register of Copyrights Barbara Ringer. She wrote:

[C]hanging any law is never an easy matter, and the case of copyright is made much more difficult by the religious fervor and theological arguments thrown at each other by the contending parties. The personal anger, the emotion, the presentation of viewpoints in stark black-and-white terms, are quite different in degree and character from what one might find in disputes over, say, admiralty or insurance law.⁵

Most of the time, the goal of Congress is to move stakeholders to a consensus solution. This approach is pragmatic, but it requires stakeholder buy in. In the context of orphan works, the copyright community was very close to consensus for several years—in part because they seemed to agree on the problem.

Consider this 2006 testimony from June Cross, a professor and documentary filmmaker. The “same problem affects all artists, big and small, publishers, libraries, museums and many others,” she said. “That is why so many of these individuals, companies and organizations have agreed that Congress must act to give protection for those users of orphan works who make a good faith effort to find the copyright holder.”⁶

Where there was tension, it was because removing statutory damages, actual damages, and injunctive relief from the arsenals of copyright owners at first seemed like very bold action, even if they were missing, asleep, or ignorant of their rights—and even though the primary goal was to pay the rights holders. The goal of connecting users to copyright owners was apparent in the Office’s 2006 *Report on Orphan Works*,⁷ and it was a concept that resonated with Congressional leadership.

For example, when introducing his bill in the 110th Congress, Senator Leahy said, “[b]y providing an incentive to search, in the form of a limitation on remedies, more users will find more owners; more works otherwise hidden will be used; and more copyright owners will receive compensation.”⁸

5. BARBARA A. RINGER, THE DEMONOLOGY OF COPYRIGHT: SECOND OF THE R.R. BOWKER MEMORIAL LECTURE 5 (R.R. Bowker Co. 1974).

6. *Orphan Works: Proposals for a Legislative Solution*, *supra* note 1, at 67 (statement of June Cross, Assistant Professor of Journalism at Columbia University and Documentary Journalist).

7. U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS (2006).

8. 154 CONG. REC. S9867-03 (daily ed. Sept. 26, 2008) (statement of Sen. Leahy).

One question that remains unresolved is the degree to which copyright owners should assert their whereabouts, if not as a question of formalities, then in proportion to the remedies they receive. Certainly, early on, some took the perspective (either on principle or out of sheer frustration) that they were entitled to full remedies under the law as a matter of constitutional interpretation and had no duty to would be users (even as a practical matter). I believe that over time this premise has shifted greatly. Big ideas have a way of bringing people along.

Many of us tend to think of the legal aspects of the orphan works debate as a two-part play: the period of 2005 through the end of 2008 and the period of 2008 to the present. The first discussion was based on an important assumption: the realization that a diligent search was necessary to reduce the number of orphan works in the sandbox, thereby keeping collateral damage to authors at a minimum and ensuring compliance with international treaties, namely the three-step test.

The second discussion was also based on an important assumption, somewhat at odds with the first one: the realization that a search that is simple to perform and relatively inexpensive, but not necessarily diligent, might attract more good faith participation among users and therefore benefit more authors and copyright owners in the long term.

Thus the question: do we need a legal framework that goes beyond the context of orphan works. In other words, do we need one that anticipates or assumes the orphans will be there, if and when we look for them.

In large part, the answer to this question depends upon whether we agree enough about the purpose and importance of mass digitization projects to provide legal comfort—a tricky question since, where public display or dissemination is the goal, mass digitization projects must be differentiated because the interests of so many stakeholders are implicated. Many of you know that Robert Kastenmeier, who was a member of the House of Representatives from 1959 to 1991, was the chairman of the Subcommittee that oversaw dramatic changes in the 1976 Act. In a lecture he gave at Columbia Law School, he summarized the Congressional thought process this way:

There are certain standards that should be applied to determine the outcome of copyright legislation. An initial standard, commonly understood by legislators, is that statutory reforms should respond

to palpable problems. . . . The proponents of change must bear the burden of proving that the change is necessary, fair and practical.⁹

This is interesting because it seems to me that many of the views about how to move forward with orphan works seem to turn on the last point, what is practical. For example, perhaps we should limit the types of works (exclude photographs for example), the types of users (museums, libraries, and archives), and the type of use (noncommercial only). Perhaps we should start small. These are worthy suggestions, and they have certainly been debated in foreign countries, but I do not love them because they do not reflect the practice of copyright.

For example, one cannot read the Office's *Report on Orphan Works* or any of the subsequent testimony without concluding that photographs are a primary point of frustration or that commercial as well as noncommercial actors require relief. And of course one often depends upon the other. We know this, obviously, from the scanning partnerships between Google and publishers and Google and libraries. But the same is true outside the context of mass digitization. Counsel for the United States Holocaust Memorial Museum testified about the many photographs, artworks, and musical compositions in the Museum's possession that require relief.¹⁰ Although the Museum is a noncommercial actor, it is not clear to everyone that museum publications are noncommercial uses within the meaning of Copyright Act, an issue that was hotly debated from 2006 to 2008. Moreover, to the extent the museum distributes books using a commercial publisher partner, as many do, the benefit of orphan works protection would likely evaporate. A solution that cannot be handed down the business chain is likely to end abruptly, short-changing the public. In comments to the Copyright Office in 2005, Duke University's Center for the Study of the Public Domain told a story of professors who were boxed in by marketplace realities. The professors said, "Penguin Classics gave us very clear limitations in terms of our publishing guidelines; since they were operating on limited budgets, there was no room to even consider any works that fell outside of 1922, even if they seemed to be free of copyright claims."¹¹

9. Robert W. Kastenmeier, *The 1989 Horace S. Manges Lecture—"Copyright in an Era of Technological Change: A Political Perspective,"* 14 COLUM.-VLA J.L. & ARTS 1, 6 (1989).

10. *Promoting the Use of Orphan Works: Balancing the Interests of Copyright Owners and Users: Hearing Before the Subcomm. on Courts, the Internet, and Intellectual Property of the H.R. Comm. on the Judiciary*, 110th Cong. 62–67 (2008) (statement of Karen C. Coe, Associate Legal Counsel, United States Holocaust Memorial Museum).

11. *Orphan Works Analysis and Proposal* at 1, Center for the Study of the Public Domain, Duke Law School (2005).

The Copyright Office heard similar stories from documentary filmmakers who cannot get insurance or distributors unless they have all rights in order, meaning all rights or the legal treatment thereof, for the partners in their business chains. Additionally, the legislation offered nonprofit institutions a special half exception or safe harbor, the option of removing works, taking them off a website, no questions asked and no payment necessary. We called this a half exception because it was conditioned on completion of a diligent search and did not allow for ongoing use without permission. However, it needs to be viewed in the larger framework and the user's determination that permission of the copyright owner is necessary in the first place and that protection against liability is worth pursuing. Otherwise, the user is free to invoke fair use.

If your heart is with the libraries and museums on this point, you may be thinking that a half exception is not especially helpful, or worth the trouble, but those who were there know that this special treatment was a real struggle and that the singular message from good faith institutions was how very risk averse they are—in other words they cannot risk fair use. Moreover, today the orphan works phrase seems to appear everywhere, from blogs to Supreme Court opinions, everyone calling for a solution. But just four years ago, Congress was inundated by complaints from artists, photographers, and a variety of American businesses, including greeting cards companies and textiles manufacturers. I know because I still have the hundreds of pieces of hate mail that came our way.

In the end, I believe that resetting the relationship between copyright owners and libraries requires both oversight and legislation from Congress, not only for orphan works but also for § 108 and to explore new forms of licensing. This is because it is in the public interest to both protect copyright owners and ensure the future of libraries. It may well be that an opt out system is the most sensible solution in some cases, but those cases require scrutiny and compromise to ensure that authors are not undercut and that library projects are responsible. These issues are a priority for the Office. Moreover, because these issues are global, as we work with Congress we are also in close touch with our counterparts in the European Union, South America, and other places.

Finally, it goes without saying that the Copyright Office must be part of the solution. The presentation and searchability of our public records should be key factors in helping copyright owners manage their rights, and users to find and assess the information. The Office should help bridge the gap between what constitutes a diligent search and one worthy of the trouble. This in fact was the focus of a group that I was part of in 1993, called ACCORD (Advisory Committee on Copyright Registration and Deposit).

Google also made this point in its orphan works comments to the Office in 2005, rather gently:

[W]e suggest that the database [of the Copyright Office] be accessible and searchable by automated means as well as through a web browser or other methods. A searchable and current database of copyrighted works would serve as a platform upon which to rest any number of legislative or regulatory improvements to the copyright system.¹²

I agree completely and have made this goal one of my key priorities. It will take funding and a multi-year strategy, as well as a commitment from Congress. The project is co-chaired by leaders from our Information Technology department, General Counsel's office, and registration program. And we will engage technical and legal experts in a number of ways before releasing our vision to Congress and the public in the next eighteen months.

In closing, I would like to again acknowledge the Principles Project for these words of wisdom:

Too much discourse about copyright law in the past fifteen years has been burdened by rhetorical excesses and an unwillingness to engage in rational discourse with those having differing perspectives. [But] it is possible for persons of good will with diverse viewpoints and economic interests to engage in thoughtful civil discourse on even the toughest and most controversial copyright issues.¹³

Thank you very much and keep the big ideas coming.

12. Google Inc., *Google's Response to Notice of Inquiry Regarding Orphan Works* 5 (Mar. 25, 2005), available at <http://www.copyright.gov/orphan/comments/OW0681-Google.pdf>.

13. Samuelson, *supra* note 2, at 1179.

PRIVATE DIGITAL LIBRARIES AND ORPHAN WORKS

Randal C. Picker[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1259
II.	THE EMERGING PRIVATE DIGITAL LIBRARIES	1262
	A. COMPETITION BETWEEN PUBLIC AND PRIVATE DIGITAL LIBRARIES	1263
	B. THE NEW PRIVATE LIBRARIES.....	1265
	1. <i>Google Book Search</i>	1266
	2. <i>The Amazon Library</i>	1268
	3. <i>Other Libraries</i>	1271
III.	COPYRIGHT RULES FOR LIBRARIES	1272
IV.	STRUCTURING THE USE OF ORPHAN WORKS	1275
	A. SELECTION EFFECTS, THE POOL OF ORPHAN WORKS, AND SECONDARY PROPERTY RIGHTS	1277
	B. SETTING A ROYALTY RATE: INCREMENTAL CREATION INCENTIVES AND ORPHAN WORKS RULES SALIENCE.....	1280
V.	CONCLUSION	1283

I. INTRODUCTION

We are in the midst of a transition to digital libraries. This process is in its earliest stages, but it is at a point where we face key structural choices that can have dramatic long-term consequences. Those choices will influence the types of libraries that will emerge and the extent of competition that we will see in the digital library space. As we build and populate digital libraries, we face critical questions about how those libraries will acquire rights to content. For current purposes, it is useful to have three types of content in mind: (1) the public domain; (2) so-called orphan works, meaning works in copyright where the copyright holder is unidentifiable; and (3) in-copyright

© 2012 Randal C. Picker. All Rights Reserved.

[†] James Parker Hall Distinguished Service Professor of Law, The University of Chicago Law School and Senior Fellow, The Computation Institute of the University of Chicago and Argonne National Laboratory. I thank the Burton and Adrienne Glazov Faculty Fund and the Microsoft Fund for their generous research support.

works where the copyright holder is known. Each of these classes poses challenges. There are no copyright limits on the use of public domain works, but there are substantial access issues nonetheless. I address those issues in a separate work.¹

Libraries face challenges in gaining access to front-list, in-copyright works—the works that populate *The New York Times* bestseller lists and the like—as the Big Six publishers assess how potential lending of library ebooks could impact sales of those works.² The problem is friction, or more precisely, the absence thereof. If library patrons can just as easily download ebooks to readers as they can purchase books, readers may abandon buying books entirely, or at least so the publishers fear. Libraries believe, of course, that they build readers and encourage reading.

This Article, however, focuses on orphan works. The problem with orphan works is well known, as is the most obvious solution: new legislation.³ Absent legislation, we are left with a world of uncertain rights and the resulting litigation, as we have seen in the Google Book Search and HathiTrust cases. But whatever mechanism emerges that allows progress on orphan works, we must attend to the opportunities available in the new landscape of digital private libraries. Current U.S. copyright law contains a number of special provisions applicable to libraries and archives and many of these rules are limited to nonprofit libraries or are limited to noncommercial activities.⁴

We would make a serious mistake if we replicate those rules for orphan works. As Part II below makes clear, we have a new emerging class of private digital libraries. This, it turns out, is an old model—dating back in the United States to Benjamin Franklin and the Library Company of Philadelphia—as private libraries were an important part of the library landscape in the eighteenth and nineteenth centuries in the United States and Great Britain. The new private digital library is therefore the return of an old business

1. Randal C. Picker, *Access and the Public Domain*, 49 SAN DIEGO L. REV. (forthcoming 2012).

2. See EBOOKS FOR LIBRARIES, <http://ebooksforlibraries.com/> (last visited Oct. 7, 2012) (demonstrating one way libraries are addressing difficulties in acquiring ebooks).

3. Some countries, such as France, are moving forward on this new legislation, often facing criticism from authors. See Gillian Spraggs, *France Guillotines Copyright*, ACTION ON AUTHORS' RIGHTS BLOG (Feb. 28, 2012), <http://blog.authorsrights.org.uk/2012/02/28/france-guillotines-copyright/>.

4. See, e.g., 17 U.S.C. § 108 (2010) (regarding limitations on exclusive rights for reproduction by libraries and archives); 17 U.S.C. § 109(b)(1)(A) (2010) (regarding rental, lease, or lending of phonorecords for nonprofit purposes by nonprofit libraries); 17 U.S.C. § 109(b)(2) (regarding lending of computer programs for nonprofit purposes by nonprofit libraries).

model that squarely offers the possibility of competition between public and private libraries and among private libraries. My central point is simple: we should not distort emerging competition by handing over special rights to orphan works to public and nonprofit libraries.

When considering orphan works, we must do so with two principles in mind. First, we need to enable broadly competing uses of the orphan works while, to the greatest extent possible, respecting the rights of the orphan works holders. Private libraries should be able to compete with each other and with public libraries in using orphan works. Of course, use of orphan works may require changes to U.S. copyright law, so it is the scope of those changes that are at stake. Second, we should not tilt the table in favor of digital library monopoly, either public or private.

The first and most critical issue for the use of the orphan works is a regime that sidesteps the current statutory damages scheme of U.S. copyright law under 17 U.S.C. § 504(c). Anyone currently contemplating using an orphan work must consider statutory damages as the great risk of that use. Actual damages are likely to be quite low, especially once we account for the selection effect at work with orphan works. Orphans, after all, are orphans for a reason. Works that are important or economically successful are not likely to take on orphan status. To revive orphan works, both public and private users must be insulated from statutory damages. And doing that will preserve competition between public and private digital libraries.

The second piece of the puzzle is, then, how we determine how much libraries should pay to use orphan works, and whether those fees should be identical for public and private libraries. Of course, if those fees are not equal, we distort the competition between the libraries. Some might favor a free-use regime or low fees for public libraries. According to this view, public libraries make works available to the public and should be supported in that mission. Free access to orphan works would do just that. The response, however, notes that a free-use regime for orphan works for public libraries imposes the cost of use on a narrow class—namely, the copyright holders for the orphan works. We could impose a good chunk of the costs of running public libraries on copyright holders, but we have refused to do that and instead generally insist that taxpayers bear those costs.

This Article is divided into three substantive parts. In Part II, this Article considers some of the characteristics that will matter for the competition between public and private digital libraries and offer a brief discussion of the emerging private digital libraries. Google's Book Search project is probably the most prominent example, but Amazon has an important initiative as well and there are many smaller examples. In Part III, this Article considers the current library exemptions in U.S. copyright law, with particular emphasis on

§ 108. Existing statutory safe harbors for libraries favor noncommercial libraries and archives. The emergence of new for-profit private digital libraries suggests that existing safe harbors reflect too narrow a conception of what libraries can be in the digital age, and we need a statutory scheme that supports the possibility of for-profit libraries. We could carry those exemptions over to private libraries quite easily and should do so to foster a rich digital library ecosystem. The possibility of large-scale private digital libraries is an important development and one we should seek to enable. If we create use rights for copyrighted works for digital libraries, we should be sure to make those privileges available to both public digital libraries and private digital libraries.

In Part IV, this Article turns to the shape of a potential statutory licensing regime for orphan works. There are good reasons to think that such a regime, whatever its setting, would play only a small role in authors' decisions on whether or not to create works. Authors will expect their works to succeed, and successful works almost never become orphans. It will be the rare potential author who will confront the possibility of orphanage for her work and then will take the next steps of assessing the likelihood that her orphan work will be restored to meaningful economic value. All of that suggests that the orphan-works licensing regime will not be particularly salient for potential authors. Even a regime that allows subsequent use at very low prices is not likely to influence the creation of very many works. Instead, we need to focus on how the licensing rules influence subsequent use of those works. In doing so, we need to account for how the performance of the orphan work will affect whether it will be claimed. Repurposed orphans that fail will find few claimants, but orphans that succeed will be claimed. This scenario is very much a heads-I-win, tails-you-lose proposition, so we will need to figure out a way to incentivize the use of the abandoned orphan works by new users.

II. THE EMERGING PRIVATE DIGITAL LIBRARIES

Access to libraries is unevenly distributed. This unequal distribution is both a product of the nature of physical libraries—tied to a particular location—and the tools we use for organizing libraries. Many of the world's great libraries are based in academic institutions. Members of those institutions often enjoy superior access to the libraries than outsiders. As for public libraries—meaning those financed predominantly through taxes—in 2011 the Public Library Association counted 9,214 public libraries in the United States, but more than forty percent of those are in towns with

populations under 5,000.⁵ These public libraries are undoubtedly important to their local communities, but the works available are of quite different scope and scale than the materials available at large metropolitan libraries, like the New York Public Library.

We are starting the transition to digital libraries, and that creates a number of opportunities. Geography instantly becomes less important, or at least can be if we make the right organizational choices. The move to digital means that new types of entities, especially private entities, may become an important part of our library ecosystem, and will likely alter the competitive balance between private and public libraries in important ways.

A. COMPETITION BETWEEN PUBLIC AND PRIVATE DIGITAL LIBRARIES

Start with public libraries and the natural question: why not move the public library online? If we finance physical libraries through general tax dollars and keep such libraries free to the public, why can we not do the same thing with online digital libraries? Some are calling this proposition the Digital Public Library of America (DPLA),⁶ but even independent of new initiatives, we are seeing our existing libraries go digital. The American Library Association (“ALA”) issues an annual report entitled “The State of America’s Libraries.” In its 2011 report, the ALA noted that ninety-four percent of all academic libraries are offering some ebooks, as are seventy-two percent of public libraries.⁷

Books are circulated as downloads or preloaded on reading devices. Actual ebook circulation is still relatively low, but is growing. The Chicago Public Library reported ebook circulation of 17,000 in 2009 and more than 36,000 in 2010.⁸ These ebooks are direct substitutes for physical books, and have not required any change in the way that libraries purchase books. The mechanics on lending and check out are a little different, but the core idea is straightforward: the library buys a certain number of digital copies, and a patron can check out the book if one of those copies is available on the virtual shelf. Again, this process assumes access to those books, and, as noted above, public libraries are facing resistance from publishers in selling ebooks.

5. PUBLIC LIBRARY ASSOCIATION, 2011 PLDS STATISTICAL REPORT 1 (2011).

6. See Robert Darnton, *Google’s Loss: The Public’s Gain*, N.Y. REV. BOOKS, Apr. 28, 2011.

7. THE AMERICAN LIBRARY ASSOCIATION, THE STATE OF AMERICA’S LIBRARIES 36 (2011), available at http://www.ala.org/news/sites/ala.org.news/files/content/mediapresscenter/americaslibraries/state_of_americas_libraries_report_2011.pdf.

8. *Id.* at 37.

Digital libraries may be at an inflection point. One path is that digital libraries will have ebooks, but the underlying information in those books will not be directly searchable. Instead, the only searchable information would be just the information card catalogues have always given us, such as author name, title, subject matter descriptors, and the like. The librarians call this “metadata.” The other path is quite different, in that libraries would act more like databases of searchable content.

For most of their lives, libraries have acted essentially as book warehouses—often very nice ones to be sure—coupled only with limited metadata designed to help users find the right books. The card catalogue was the traditional way in which this metadata was presented to library users. For libraries, the initial move to digital was “retrospective conversion,” which describes the process for taking card catalogue metadata online, allowing patrons to search metadata directly. But truly digital libraries offer so much more: the ability to perform full-text searches on the entire corpus of their content, just as we do with Google and Bing and other search engines every day.

Return to the three types of works classified above—the public domain, orphan works, and active in-copyright works—and consider competition between public and private libraries. The public domain is open to all, though this public availability presents issues more interesting than that simple formulation suggests.⁹ This Article is focused on orphan works and their influence on the digital library market, but in-copyright works, or commercial works, will also play an important role in the development of digital libraries.

Our online public libraries may struggle to get access to these commercial works. Copyright holders will be nervous that ease of use of a digital public library will cause consumers to purchase fewer books. Although public library ebooks are still at an early stage, publishers are already adjusting how they approach ebook sales to libraries. In a move that generated widespread discussion in February of 2011, HarperCollins announced that, going forward, its library ebooks would expire after twenty-six check outs.¹⁰

After all, physical books degrade, but digital books have a much longer natural lifespan, so although publishers might avoid the costs of producing physical copies by transitioning to digital, they may also lose revenue from re-sales. The Google Book Search settlement called for a split of the revenues that the project would create—roughly thirty-seven percent to Google and

9. That was cryptic and, I hope, a teaser. For more, see Picker, *supra* note 1.

10. See Calvin Read, *Librarian Unhappiness Over New Harper e-Book Lending Policy Grows*, PUBLISHERS WKLY., Mar. 2, 2011.

sixty-three percent to rights holders.¹¹ A free online public library version of a book database would not generate revenues of this sort. It seems unlikely that copyright holders would be satisfied selling one never-checked-out, searchable copy of a digital book to the public digital library for \$14.99, or whatever ebooks go for these days, if that book would then be provided free of charge digitally.

As the above example suggests, the contracting process for in-copyright works with active rights holders will not be simple. Copyright holders will be looking for revenue streams and will have the full right to prefer revenue-generating private digital libraries over a free online public library. Many electronic databases are sold today with lump-sum payments, but that was not the GBS model, and we should be skeptical that copyright holders will accept a one-time, lump-sum access fee model.

Instead, it is more likely that digital public libraries will track the GBS deal by paying on a usage basis. The 37/63 revenue split in the Google settlement was a usage deal tied to revenues. We could imagine public libraries that charged for the use of books—financed through user fees (including the possibility of price discrimination) rather than through general tax revenues—but that would be controversial and, in any event, the public library would probably be operated on a non-profit basis. Revenues are not likely to be the basis for usage fees paid by public libraries to copyright holders. Instead, we might expect deals that tracked use, though of course such tracking will require us to quantify use and to deal with the equivalent of click fraud.

We are likely to see competing approaches to financing, privacy, scanning, metadata, and search. Public libraries may face restrictions on their ability to track users, though these restrictions will also serve as a competitive benefit as public libraries can promise greater privacy to their users. At the same time, as discussed below, private libraries may be able to draw on new revenue streams unavailable to public libraries.

B. THE NEW PRIVATE LIBRARIES

As detailed below, we are seeing the emergence of a new class of private digital libraries that hold, perhaps ironically, the promise of greater equality in access. Access to physical public libraries is limited by distance. Online private libraries may be accessible anywhere there is an Internet connection and a device, such as a smartphone. To be sure, these private digital libraries

11. See Isabel Howe, *How the Settlement Will Work*, THE AUTHOR'S GUILD, <http://www.authorsguild.org/advocacy/articles/settlement-resources.attachment/how-the-settlement-will/How%20the%20Settlement%20Will%20Work.pdf>.

will be tied to a digital infrastructure, which does bring its own limits, though linking these libraries with our physical public libraries can help to overcome some of those limitations. Furthermore, digital access costs have dropped rapidly over time, consistent with Moore's law,¹² while the costs of paper, brick, and mortar are not changing rapidly.

These private libraries also represent a return to an older library model. There is a substantial history of private libraries in the United States and the United Kingdom. Unsurprisingly, there is an extensive nomenclature for classifying libraries,¹³ but my focus here is on libraries neither affiliated with universities or churches, nor funded by the government. These encompass social libraries, as they are known, where members band together to fund access to books made available to the membership or even to the public. In the United States, Benjamin Franklin is credited with launching the first library of this sort in 1731,¹⁴ and there were corresponding and perhaps even slightly earlier developments in the U.K.¹⁵

1. *Google Book Search*

At the 2004 Frankfurt book fair, Google announced its new Google Print service, called Google Book Search ("GBS").¹⁶ Consistent with its corporate mission statement—"to organize the world's information and make it universally accessible and useful"¹⁷—Google turned its attention to the great treasure trove, the information collected by the world's great libraries. This was a bold and complicated undertaking,¹⁸ but the short version of the story is that when Google set out to digitize the contents of these libraries, it stepped into copyright quicksand. Google almost certainly knew what it was doing, so it was not likely surprised by the resulting lawsuits, litigation, and

12. Moore's law describes the long-term trend that the number of transistors that can be placed inexpensively on an integrated circuit doubles approximately every two years, resulting in a roughly exponential increase in the capabilities of many digital devices. Gordon E. Moore, *Cramming More Components onto Integrated Circuits*, 38 *ELECTRONICS*, no. 8, Apr. 19, 1965, available at http://download.intel.com/museum/Moores_Law/Articles-Press_releases/Gordon_Moore_1965_Article.pdf.

13. See, e.g., HAYNES McMULLEN, *AMERICAN LIBRARIES BEFORE 1876* (2000).

14. *Id.* at 22.

15. See Richard Roehl & Hal R. Varian, *Circulating Libraries and Video Rental Stores*, *FIRST MONDAY.ORG* (May 7, 2001), <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/854/763>.

16. See Edward Wyatt, *New Google Service May Strain Old Ties in Bookselling*, *N.Y. TIMES*, Oct. 8, 2004, available at <http://www.nytimes.com/2004/10/08/technology/08book.html>.

17. ABOUT GOOGLE, <http://www.google.com/about/company/> (last visited Oct. 7, 2012).

18. I describe Google's effort in detail in Randal C. Picker, *The Google Book Search Settlement: A New Orphan-Works Monopoly?*, 5 *J. COMPETITION LAW & ECON.* 383 (2009).

eventual settlement. Whether it also anticipated the rejection of that settlement is another question, and the litigation is ongoing.¹⁹

It is hard to know how Google believed its library would operate. As launched, the project was a standard—for Google at least—two-sided undertaking: content was free to consumers but advertisers would pay to reach those consumers. That arrangement is a striking departure from the way in which libraries have been financed traditionally. Think of this as library-as-loss-leader or as commercially-subsidized reading. Google was providing books to consumers below the cost of creating the service, but was relying on advertisers to pay the bills. And the fact that Google's library was situated as part of its overall business meant that the GBS's economics were fundamentally different than those typically associated with libraries. Google clearly hoped that the book digitization would reverberate to the benefit of their other businesses, such as search and its translation service.²⁰ By operating in multiple lines of businesses, Google can take advantage of scale and scope economies. Google's search tools improve as they are used in each of their business lines and, GBS readers would get the advantage of Google's general search improvements, plus their access would be subsidized by the advertisers eager to reach them.

The original version of the service offered general access to the digital public independent of location. The world of physical libraries is unevenly distributed, but GBS would be available to anyone with a digital device and a broadband connection. Individuals without those tools would be able to participate through physical public libraries with access to the service.

19. *Authors Guild v. Google, Inc.*, 770 F. Supp. 2d 666 (S.D.N.Y. 2011). While the lawsuit with the authors continued, Google and the publishers took another stab at a settlement in early October 2012. See Claire Cain Miller, *Google Deal Gives Publishers a Choice: Digitize or Not*, N.Y. TIMES, Oct. 4, 2012, available at <http://www.nytimes.com/2012/10/05/technology/google-and-publishers-settle-over-digital-books.html>.

20. See GoogleTechTalks, 2011 Frontiers of Engineering: Research at Google Lightning Talks, YOUTUBE (Sept. 28, 2011), <http://www.youtube.com/watch?v=YGBZGhzLL14> (presenting National Academy of Engineering, U.S. Frontiers of Engineering Symposium, September 19–21, 2011) (see the discussion at 3–6 minutes); see also *Objection of Yahoo! Inc. to Final Approval of the Proposed Class Action Settlement at 25, Authors Guild, Inc. v. Google Inc.*, No. 05 CV 8136 (DC), 2009 WL 2980749 (S.D.N.Y. Sept. 8, 2009) (stating that Google Book Search will secure for Google “access to a database capable of vastly improving Google’s central product, Google Search”); Jeffrey Toobin, *Google’s Moon Shot*, NEW YORKER (Feb. 5, 2007) http://www.newyorker.com/reporting/2007/02/05/070205fa_fact_toobin?currentPage=4 (explaining that the database includes a primitive translation feature that “may someday be enhanced to allow books to be rendered in another language at the touch of a button”).

And this would be broad access. Most public libraries outside of large cities in the United States offer only limited access compared to what Google would offer through GBS. Google stocked its library by entering into partnership agreements with many of the world's great libraries. Google could digitize the public domain at will and put those books online and into the hands of readers everywhere. The genius of the proposed settlement to the Author Guild's class action litigation against Google was that it promised to enable wide access to the orphan works. And Google would get whatever rights to actively-managed in-copyright works that it could obtain through negotiation with rights holders.

With the rejection of the proposed settlement, the GBS project faces an uncertain future. There are the usual possible damages issues for alleged past acts of copyright infringement, although Google, of course, will assert fair use and other possible defenses. As for the service itself, Google can move forward on offering access to works for which it has negotiated the rights to, and for public domain works. It is only the orphans that may be left out.

2. *The Amazon Library*

On November 2, 2011, Amazon announced a new Kindle Owners' Lending Library, drawing a number of trends sharply into focus. Like GBS, this was another private effort by a key player in the digital space. Unlike GBS, this effort offered immediate zero (marginal) cost full-text access to in-copyright works to many Amazon customers. Prior to the launch of Amazon's library, it was rumored that Amazon would launch a Netflix for books.²¹ The original Netflix model was to provide a fixed number of DVDs to customers at a time, with a new DVD mailed out as soon as another was returned. But Amazon's library is actually much more limited than that: you can check out no more than one book at a time and only one book each month. Yes, twelve books for an entire year. But the service would be free to Amazon Prime customers.

That seems like a library that only readers who linger over their books could love, so why does Amazon impose such severe borrowing limits? To answer this question, we need to look at how Amazon gets books into its library and how it pays for those books. A library, of course, needs books. Amazon had approached the Big Six publishers—Random House, Simon & Schuster, HarperCollins, Macmillan, Penguin, and Hachette—to try to reach

21. Matthew Ingram, *Why the Idea of a Netflix for E-books Makes Sense*, GIGAOM (Sept. 12, 2011, 3:04 PM), <http://gigaom.com/2011/09/12/why-the-idea-of-a-netflix-for-e-books-makes-sense/>.

a deal with each to include books they published in the library, but each declined to participate.²²

Book publishers have long been skeptical about libraries, so it is hardly surprising that they would feel especially so about a new library type. The basic fear, as articulated by Houghton Mifflin Harcourt regarding Amazon, is that “we do not want lending to replace selling.”²³ If books are sold for a set price to all purchasers, efficient sharing of books could lead to reduced sales. Books can be shared by making used book markets efficient, and book industry insiders have long been unhappy with Amazon for facilitating used book sales.²⁴ But book sharing is exactly why libraries exist, and a library run by a for-profit entity like Amazon might lead to especially efficient book-sharing.

But shuffling around lots of physical books would be expensive. Not even Netflix really wants to be Netflix anymore, as its aborted plan to hive off its DVD business as Qwikster made clear.²⁵ Amazon is driving the transition to ebooks with its Kindle platform, so an Amazon library would be digital from birth, making content acquisition much more complex. You buy physical books and can do with them as you please, but Amazon has contracts to get ebooks, and those contracts might impose limits on which books Amazon could add to its library. That seems to be the case, though it is hotly contested as to what exactly the contracts permit as between Amazon and the publishers and between publishers and authors.²⁶

Amazon announced its new program as having “thousands” of books to borrow, including more than one hundred current and former New York

22. Calvin Reid, *Amazon Launches Lending Library without the Big Six*, PUBLISHERS WKLY., Nov. 2 2011, available at <http://www.publishersweekly.com/pw/by-topic/digital/content-and-e-books/article/49374-amazon-launches-lending-library-without-the-big-six.html>.

23. *Id.*

24. David D. Kirkpatrick, *Online Sales of Used Books Draw Protest*, N.Y. TIMES, Apr. 10, 2002, available at <http://www.nytimes.com/2002/04/10/technology/10BOOK.html>.

25. Peter Kafka, *Qwikster is Gonester: Netflix Kills Its DVD-Only Business Before Launch*, ALL THINGS D (Oct. 10, 2011, 5:01 AM), <http://allthingsd.com/20111010/qwikster-is-gonester-netflix-kills-its-dvd-only-business-before-launch/>.

26. See *Contracts on Fire: Amazon's Lending Library Mess*, AUTHORS GUILD BLOG (Nov. 14, 2011), <http://blog.authorsguild.org/2011/11/14/contracts-on-fire-amazons-lending-library-mess-2/>; Simon Lipskar, *Author Contracts and Subscription Models*, AARDVARK (Nov. 4, 2011), <http://aardvarknow.us/2011/11/04/author-contracts-and-subscription-models/>; Jennifer Wertz, *The Biggest Unasked Question With the New Kindle Lending Library—How are Authors Paid*, AARDVARK (Nov. 4, 2011), <http://aardvarknow.us/2011/11/04/the-biggest-unasked-question-with-the-new-kindle-lending-library-how-are-authors-paid/>.

Times bestsellers.²⁷ Amazon's press release indicated that for "the vast majority" of titles, Amazon had reached a deal with publishers to include their books for a fixed fee.²⁸ In other cases, Amazon was buying a copy of the ebook at a wholesale price each time the book was borrowed. This practice is the opposite of efficient book sharing. In an effort to demonstrate the value of the library, Amazon was buying books at wholesale and giving them away to Amazon Prime customers. Even in doing this, some publishers contended that Amazon had overstepped the publisher contracts with Amazon, while authors claimed the same about publishers.²⁹

Given the chilly reception from the Big Six publishers and leading organizations of authors, Amazon found friendlier territory with its Kindle Direct Publishing ("KDP") authors. On December 8, 2011, Amazon announced a new KDP Select program to add books to the Kindle Library.³⁰ KDP authors would agree to commit their books to Amazon exclusively for at least ninety days. Amazon, in turn, would include those authors' books in the Amazon library. To induce authors, Amazon offered a \$6 million carrot to be split by authors based upon library downloads. Authors would split the fund based on their pro rata share of library downloads.

Just to frame the \$6 million using a parochial example, I grew up in Akron, Ohio and currently live in Chicago. Akron is now the 110th largest city in the country and the entire acquisition budget for the Akron-Summit County Public Library was just about \$3 million in 2010 and \$3.4 million in 2011.³¹ For Chicago—the third largest city in the United States—the 2010 acquisition budget for books and materials was \$17.8 million.³² And the \$6 million from Amazon is just for the KDP Select program; it does not include what Amazon is spending on non-Amazon-based authors.

27. Press Release, Amazon.com, Introducing the Kindle Owners' Lending Library (Nov. 2, 2011), *available at* <http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle&ID=1625426>.

28. *Id.*

29. *See Contracts on Fire*, *supra* note 26 (stating that Amazon got away with including books in the program by "giving its boilerplate contract with these publishers a tortured reading").

30. Press Release, Amazon.com, Introducing "KDP Select"—A \$6 Million Fund for Kindle Direct Publishing Authors and Publishers (Dec. 8, 2011), *available at* http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle_Print&ID=1637803&highlight=.

31. AKRON-SUMMIT COUNTY PUBLIC LIBRARY, 2011 PERMANENT BUDGET: GENERAL FUND (2011), *available at* <http://www.akronlibrary.org/2011Budget.pdf>.

32. CHICAGO PUBLIC LIBRARY & CHICAGO PUBLIC LIBRARY FOUNDATION, 2010 ANNUAL REPORT 12 (2010), *available at* http://www.chipublib.org/dir_documents/ar_2010.pdf.

How is Amazon making money from this? Amazon's business ecosystem is linked together tightly. The Kindle library program is limited to actual Kindles, that is, Kindle devices sold by Amazon. The Kindle platform is broader than that with apps for Apple's iOS, for Android, and for computers. But those Amazon users are not eligible to borrow books from Amazon's library. And the library is part of Amazon's Prime program. The Amazon prime program started as a shipping program—pay a \$79 annual fee and receive “free” two-day shipping on your Amazon purchases—but now has expanded into a bundle of shipping, video access, and the Kindle library.

3. *Other Libraries*

The Google and Amazon projects are the most visible of the new private digital libraries, but we are likely to see more modest efforts as well. Take just one example: on March 29, 2012, F+W Media announced what it described as the first in a planned series of ebook subscriptions.³³ The first site offered more than one hundred full-color art instruction titles for an annual subscription fee of \$199. Artists are not able to download the books, but can instead access the books as needed so long as their subscriptions are active.³⁴

The Artist's Networks eBooks Art Book Club performs one of the services we expect from libraries. Book ownership entails instant and unlimited access to a book. Typically, you own books when you do not want to share them with others. But there are some books where occasional access suffices. You do not need exclusive access; you just want to consult these books as needed. This is the exact scenario the Art Book Club seeks to address by providing a vertical library organized around a particular subject area.

We can see the advantages of the digital age for this type of content sharing. Physical libraries must be reasonably close to their users. We may be able to share pieces of plastic at a distance—this type of sharing is what Netflix does after all, though, as noted, even Netflix wants to drop physical media for digital distribution—but the bigger and heavier the media, the less feasible sharing by mail becomes. Vertical, specialized libraries will face natural limits when they are dealing in atoms, but switch to bits, and assuming that the works can be shared securely, you have a business like the Art Book Club.

33. See Press Release, F+W Media, Inc., F+W Media Releases First eBook Subscription Site for Enthusiasts (Mar. 29, 2012), available at <http://www.fwmedia.com/press-room/fw-media-releases-first-ebook-subscription-site-for-enthusiasts>.

34. See *Frequently Asked Questions*, ARTISTSNETWORK: EBOOKS, <http://ebooks.artistsnetwork.com/faqs> (last visited Oct. 9, 2012).

These are exciting developments in new libraries. The emergence of new libraries that are open to the public at large and not tied to the happenstance of location or university affiliation is fundamentally democratic. To be sure, you need money to participate in Amazon, but it allows you cross the digital divide at a reasonable price. For \$149, a user may purchase a Kindle Touch 3G and receive access to the Kindle library, plus large numbers of public domain works for no additional price. Vertical libraries will require money as well, but one can imagine libraries financed through advertising, as Google planned, in part, with Google Book Search.³⁵

III. COPYRIGHT RULES FOR LIBRARIES

Libraries could change rapidly, but copyright law could either slow down or accelerate the pace of change. Both public and private digital libraries can offer the public domain, though I do think that distributing the public domain is more complex than often recognized.³⁶ For active works in copyright, libraries will need to acquire rights to those works via direct consent, absent some sort of statutory compulsory license. But my interest here is in the orphan works where we could imagine special copyright legislation that would facilitate certain uses. Indeed, in January of 2006 the U.S. Copyright Office issued an extensive report suggesting how to modify copyright law to thus enable wide use of orphan works.³⁷ In 2008, proposed legislation made it through the Senate, but the legislation died in the House.³⁸

With the reports and proposed legislation in mind, we should assess how U.S. copyright law approaches libraries and archives. I am never quite sure whether U.S. copyright law is a general framework with many situation-specific special rules and cases or whether it is really just all special rules. Perhaps that exaggerates slightly, but there is no question that while you need to have a good sense of the overall approach of copyright law, you also need to pay attention to exactly where your case is situated. When looking at libraries and archives, we should consider why we might have special rules for those institutions and what these rules might look like. Libraries are an exercise in sharing; public libraries are an exercise in ensuring widespread and democratic access to works. Libraries and archives are also the institutions

35. See Wyatt, *supra* note 16.

36. See Picker, *supra* note 1.

37. See U.S. REGISTER OF COPYRIGHTS, REPORT ON ORPHAN WORKS (2006), available at <http://www.copyright.gov/orphan/orphan-report.pdf>.

38. See *Bill Summary & Status for S.2913, the Shawn Bentley Orphan Works Act of 2008, 110th Cong.*, LIBRARY OF CONGRESS, <http://www.thomas.gov/cgi-bin/bdquery/z?d110:SN02913:@@@R> (last visited Oct. 9, 2012).

we count on to preserve works and ensure continuity over time in the objects in which the works are embodied. The organization of this sharing turns on the needs of users, the costs of the objects being shared, and the costs of sharing.

We could imagine how these roles for libraries might show up in exemptions from the core rights of copyright holders set forth in § 106.³⁹ If we allowed public libraries to copy books fully without fear of liability for copyright infringement, we would maximize widespread access. We could also imagine that a regime of free copying would ensure preservation of works, as libraries would copy works and disseminate them to other libraries.

Of course, a regime where a copyright holder could only count on *one* sale to libraries, as a group, would require that the sale be made at a very high price, or that authors allow their books to be widely available with very little revenue in return. That is not the world that we live in, at least in the United States. Although there are many special provisions for libraries and archives in U.S. copyright law, I can best frame the discussion by focusing on § 108. Section 108 addresses a number of particularized situations, but what I think we will see is that we do not use these rules to subsidize libraries. We, of course, subsidize public libraries in the most basic sense possible—tax revenues—but we have on the whole avoided creating special within-copyright subsidies that would be borne by only a narrow slice of the population, namely, copyright holders. We subsidize public libraries with general taxes rather than narrow taxes targeting copyright owners.

For example, consider the rules for preserving unpublished works in § 108(b).⁴⁰ A library or archive holds an unpublished, deteriorating work. Since the work is unpublished, you cannot go into the market to buy a replacement copy. Given that a buyer cannot buy a replacement copy of an unpublished work, we should not fear that copying will substitute for a potential purchase. Moreover, absent the copying, this particular library or archive will lose ready access to the work and such a loss may reduce overall access as well, as would be the case if there were few copies of the unpublished work or this particular copy was the last one. This seems like socially beneficial copying.

Nothing in the analysis turns on how the library is financed. The library could be a public library, financed primarily through taxes. It could be a nonprofit library such as a college or university library. But it could also be a for-profit library organized by Google, Amazon, or some new entrant. The

39. 17 U.S.C. § 106 (2010).

40. *Id.* § 108.

point of the preservation exemption is just that—preservation—and it is hard to see how the social interest in preservation drops merely because there might be money on the table somewhere.

Section 108(b) sets out a rule for unpublished works. There is a corresponding rule for published works in § 108(c). A key aspect of this rule requires that the library or archive be unable to obtain an “unused replacement” at a “fair price.” Again, this rule reflects the core substitution idea that drives copyright: we try to control users’ efforts to substitute out of purchases of new copies of a work. If you operate a library and destroy a new physical copy of the latest bestseller in a vat of preservation glue, you need to go buy another copy of the book. There is no general built-in insurance scheme against this kind of destruction. But there is a more limited type of insurance made available: if you cannot buy a new copy readily—and the fair price provision is clearly intended to limit unusual situations of market power—the library or archive can make a copy, subject to a few more restrictions.

Again, the point of § 108(c) is to preserve works in circumstances where normal sales are not displaced and nothing seems to turn on the financing or profit status of the library. A library can either buy the book at a fair price or it cannot. Copyright holders do not want libraries exiting purchases in circumstances where they can buy on standard market-place terms and the statute does not allow libraries to do that.

There are a couple of possible, subtle responses to this issue. One response is that we should limit these preservation exemptions to public and nonprofit libraries in an effort to drive rare works into public hands. The notion is that a private library holding a deteriorating work would recognize that it could preserve it only by giving it up to a public institution. If you love it, let it go. But this idea seems optimistic and puts works at risk if private owners hold onto works for a little too long.

Another possible response is that libraries underinvest on the front end in multiple copies of a work given that they recognize that they can just make more money at the back end. That idea probably matters most for published works as the unpublished works probably have never been available for per-copy purchase. Of course, the most successful works will have remained in commercial exploitation and, if their current copies wither, libraries will have to buy replacement copies of those works. For works that the library expects to succeed and that actually turn out to be successful, libraries cannot opportunistically under buy. Whatever we think of this idea—and it seems a little speculative—there does not seem to be any obvious reason why we should allow public libraries to under buy while denying the same right to private library competitors.

Recall what we are doing here: we are looking at § 108 to see what we can glean regarding how U.S. law approaches permitted copying by libraries. The two § 108 exemptions considered are tied to preservation. Two other exemptions contemplate a different situation, namely ones where users of libraries or archives seek copies for “private study, scholarship or research.”⁴¹ These exemptions focus on what the recipient will do with the copy. Section 108(d) addresses copying from periodicals and similar works, while § 108(e) addresses copying an entire copyrighted work. The entire-work exemption is again limited by the condition that “a copy or phonorecord cannot be obtained at a fair price.”⁴²

This high-level summary of § 108 emphasizes that the basic frame of § 108 is not about subsidizing libraries. To be sure, the conversation about subsidizing libraries turns a little on how you see the over-time substitution questions, meaning the extent to which the § 108 exemptions enable libraries to buy less upfront given that they know that they can engage in some ex post copying. Do not buy two copies of a work now given that you can make another copy later if the book is no longer commercially available or do not subscribe to a journal given that you know that your patrons will be able to get copies made from other libraries.

Obviously, we do subsidize public libraries when we pay for library operations through general tax revenues, but, as the discussion of § 108 demonstrates, we do not subsidize public libraries through some sort of special within-copyright, in-kind subsidies. It would be easier to run public libraries if they received free paper and pens, yet we do not require Office Depot to ship supplies to the libraries for free. It would also be easier to have public libraries if they received free books, which we have not done, or if we granted public and noncommercial libraries free access to the orphan works. That means that, as we consider an orphan works licensing regime, we should focus on the extent to which we should depart from that basic framework in constructing an orphan works licensing regime for both public and private libraries.

IV. STRUCTURING THE USE OF ORPHAN WORKS

We can leave the orphan works outside of our new digital libraries, but that would be a substantial social loss and almost certainly a loss to the missing copyright holders who hold the copyright to those works. The core premise of the activity surrounding orphan works is that we should put them

41. *Id.* § 108(d)–(e).

42. *Id.* § 108(e).

to use if we can come up with a sensible structure for doing so. The animating principle of such legislation should try to replicate what we think orphan rights holders would do were they actually present. Legislation that captured this idea would exhibit the greatest fidelity to the existing copyright system. Orphan-works legislation should not be seen as an opportunity for giving orphan holders weaker rights merely because they are not present and are unrepresented. For example, copyright holders do not typically give books for free to public libraries. That said, we do need to take into account the selection effects associated with orphan works—more on that momentarily—when framing possible royalty rates for orphans. If the royalty rate was set at rates similar to those for comparable present rights holders, there are good reasons to think that we would be overpaying. Should orphan holders never come forward, those payments might be escrowed pending identification of owners of the orphan works and those escrow payments might escheat back to the government under normal rules.

We are clearly at an early stage in our transition to digital libraries. All of that suggests that we should expand our conception of what a library is and that we should not tilt the legal tables in favor of public or private libraries. Only the government can create a license for the orphan works and I am hard-pressed as a matter of first principles to understand why that license should be limited. Such a license should not run in favor of one party nor should it be limited, as suggested by Robert Darnton, to entities that wish to make noncommercial uses of those works.⁴³ New orphan-works legislation should enable broad competing uses of the orphan works, by both commercial entities and non-profits.

This concept should be explored in greater detail. There seem to be four natural rules regarding the use of orphan works, what I will refer to as the orphan works licensing regime: (1) create no special rules regarding exempted uses; (2) allow use for free, which is to treat the orphan works as if they have entered the public domain; (3) allow use and insist that all users pay for that use on the same scale; and (4) allow uses and create a payments menu where commercial and non-commercial users pay different amounts (and where non-commercial users might pay nothing). It is important to keep in mind that the orphan works regime will probably operate within the broader set of rules for exempted uses, including fair use. As such, the design question for orphan works is also very much a rules/standards question.

43. Robert Darnton, *A Digital Library Better than Google's*, N.Y. TIMES, Mar. 24, 2011, at A31, available at http://www.nytimes.com/2011/03/24/opinion/24darnton.html?_r=0.

Fair use has the great virtue of flexibility and that means fair use can evolve as circumstances warrant without the need for further legislation. The flip side of this outcome, of course, is that fair use is also subject to litigation risk. Exempt uses, such as those set forth for libraries in § 108, are classic rules: reasonably well-defined inflexible straightjackets with minimal risk. Part of what we need to consider when designing an orphan works regime is how much of the framework should be driven into a rules-like system and how much of it can be left to more open-ended standards.

Much of the actual complexity that we see in proposed orphan works rules arises from the classification question of how we identify an orphan work, or, same point put differently, the entry and exit rules for orphan works. How is a particular user supposed to establish whether a particular work is actually an orphan? What happens if they get this determination wrong? What happens if something changes and a work that was once an orphan ceases to be orphaned? The answers to these questions are obviously important in defining how an orphans works regime will work in practice, but such questions are not the focus of this Article. Instead, I want to focus on the royalty rate questions and the issue of whether we need some form of secondary right property regime for subsequent users of the orphan works.

A. SELECTION EFFECTS, THE POOL OF ORPHAN WORKS, AND
SECONDARY PROPERTY RIGHTS

Consider the first possible rule, which is basically current law: no special exemptions tied to orphan status.⁴⁴ This rule means that using an orphan work brings with it the risk of statutory penalties. Obviously, the statutory penalties regime applies to all works in copyright, so the natural question is whether a different rule should be created for the orphan works. Part of this has to be driven by a sense of the social waste that arises when orphans are excluded from use because of that status. Note, however, that this is not a general statement about non-use. We know that there are some copyright holders who want to restrict distribution of at least some of their works, such as their private letters. J.D. Salinger comes to mind.⁴⁵ For these authors, the best way to control their works is to manage them quite carefully, and therefore those works are unlikely to become orphans.

44. In the current statute, 17 U.S.C. § 108(h) is described as a limited orphan works right in that it creates a special distribution right for libraries and archives during the last twenty years of any term of copyright, but not if the work is subject to commercial exploitation or the copyright owner certifies that it is being so exploited.

45. *Salinger v. Random House, Inc.*, 811 F.2d 90 (2d Cir. 1987) (deciding that J.D. Salinger could prohibit both the quoting and paraphrasing of his unpublished letters).

Instead, orphans arise when the expected benefits of controlling copyright to a work are exceeded by the costs of maintaining a title system to the work. There is a strong selection effect at work as to the works that become orphans. Works that are making money year-by-year should not become orphans as there are good incentives for maintaining the chain of title. We should expect chain of title weaknesses— orphan status—to arise for the books that have not sustained themselves economically or in cases where copyright holders are otherwise uninterested in controlling their works.

In the past, books have exited commercial status because of distribution costs associated with physical books. Books were not sold through a central point, such as Amazon.com, but instead were sold through distributed bookstores. Shelf space in those stores is finite, so we would expect low-selling books to be crowded out by books selling more copies. A book that fell beneath the shelf-space cutoff would exit distribution and then, subsequently, exit print. Contracts between authors and publishers typically provided for copyright to revert to the author at some point when the book exited print.

We do not live in that world today. Instead, we live in a world of eBooks and the possibility of print-on-demand and centralized distribution. That possibility means that fewer books will leave print and presumably we will see fewer orphans arise. Orphans have arisen because of the death of authors and the death of publishers. Fewer reversion means that author death will become less important as a source of orphan works. Contracts between authors and publishers are changing to reflect this situation, with many contracts giving authors weaker reversion rights.⁴⁶

All of that suggests that the pool of orphan works will grow more slowly going forward. But, at present, we have a stock of existing orphan works and we need to figure how to manage those works. Today, our orphan works are very much the abandoned, broken toys at the back of the closet in the *Toy Story* movies. We want to create incentives for people to find those toys, pick them up, and polish them. Indeed, we might want to create a regime to incentivize discovery and use of abandoned works. Suppose that a publisher discovers an out-of-print work that could be revived given recent trends in book readership. This is happening today as we are seeing out-of-print books

46. Hillel Italie, *In print forever? Sometimes, authors would rather go out of print*, SEATTLE TIMES, May 21 2007, available at http://seattletimes.nwsourc.com/html/entertainment/2003714078_webprint21.html.

revived as eBooks and then also in print as print-on-demand books.⁴⁷ Assume that the publisher conducts an appropriate rights search to determine who holds copyright to the work but is unable to track down the owner.

In that scenario, we should fear that the book will sit dormant. Absent an orphan works regime, the publisher could face substantial penalties for copyright infringement. A use regime for orphan works would control this exposure. But there is another issue as well. Users who seek to revive orphan works face the reality that they will be stuck with the failures while market success will be sure to attract attention and likely claims of ownership. We do not really have a copyright doctrine of abandonment or an IP analog to adverse possession, such that we create a regime of secondary or follow-on rights in the reviver of a particular orphan work.

If we were to embrace a quasi-abandonment posture for the orphan works, we could imagine a number of possible results. Some might favor treating the orphans as if they had entered the public domain, meaning, operationally, no monetary remedies for the use of the orphaned work. Recall that the 2006 report of the U.S. Copyright office favored exactly that result, at least for noncommercial users.⁴⁸ A preference for a public domain approach to abandoned orphans turns somewhat on whether you think that more limited property rights are required to get orphans put back to use in the first place. We have done something like that in the Hatch-Waxman scheme for encouraging entry by generic drug manufacturers at the expiration of a prior patent.⁴⁹ Allowing a reviver to continue to use the work in the face of a competing claim by a now-appearing parent would undercut a key part of copyright, but we might need to accept that sort of change to help restore orphans to use.

47. Alison Flood, *New Bloomsbury digital imprint revives hundreds of neglected classics*, GUARDIAN, Sept. 28, 2011, <http://www.guardian.co.uk/books/2011/sep/28/new-bloomsbury-digital-imprint>.

48. *See supra* note 37.

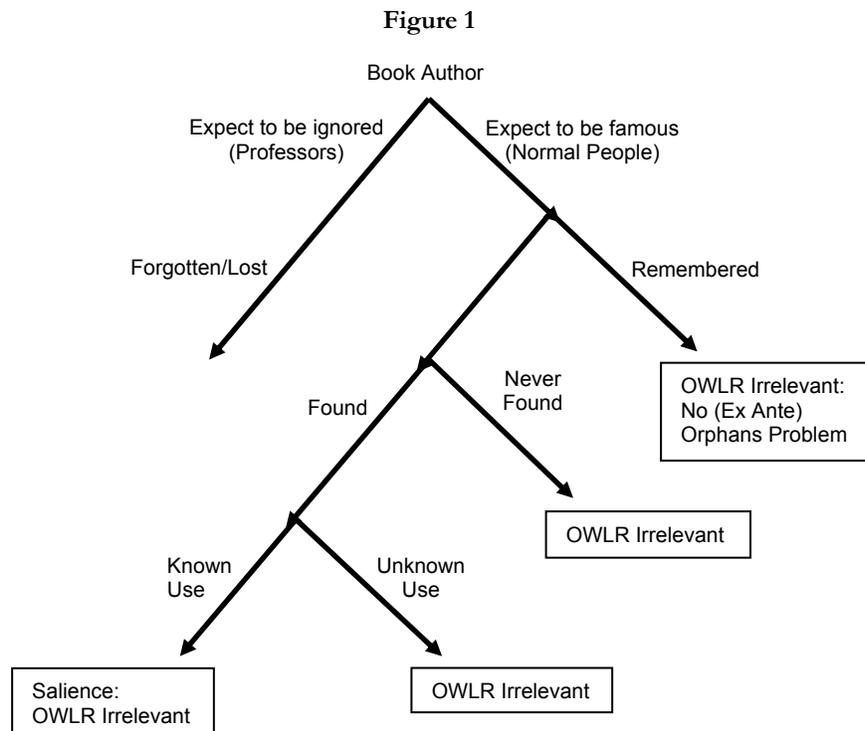
49. The Hatch-Waxman scheme comes from the Drug Price Competition and Patent Term Restoration Act of 1984 Pub. L. No. 98-417. At the conference associated with this paper, Lydia Loren offered another approach to reclaiming abandoned works—hostage works in her terms. *See* Lydia Loren, *Abandoning the Orphans: An Open Access Approach to Hostage Works*, 27 BERKELEY TECH. L.J. 1431 (2012).

B. SETTING A ROYALTY RATE: INCREMENTAL CREATION INCENTIVES
AND ORPHAN WORKS RULES SALIENCE

If we move forward with special use rules for orphans, as has occurred in France⁵⁰ already and is being considered more generally in the European Union, we will have to confront the critical question of royalty rates. I suggested three possibilities in the earlier discussion: (1) free access via the public domain; (2) paid access with the same royalty rate for all users; and (3) paid access with rates tracking status. As I have suggested, we do not generally expect copyright holders to subsidize libraries, and a regime of free use of orphan works seems to be exactly that sort of subsidy. This reality obviously reveals a broad social question about the interaction of general taxes, the public interest, copyright, libraries and the scope of the public domain and I will not take it on here in this Article.⁵¹ Instead, I will assume a positive rate and turn to considering the question of how we should think about royalty rates for the use of orphan works.

50. See Assemblée Nationale, Texte Adopté No. 862 (Feb. 22, 2012), available at <http://www.assemblee-nationale.fr/13/ta/ta0865.asp>; Spraggs, *supra* note 3.

51. The U.S. Copyright Office's Report on Orphan Works proposed that noncommercial libraries face no monetary penalties for the use of orphan works. See REPORT ON ORPHAN WORKS, *supra* note 37, at 127 (proposing a new § 514 creating limitations on remedies in connection with the use of an orphan work).



I want to start by focusing on the role that an orphan works licensing regime would likely play in inducing the creation in the first place of copyrighted works. Consider a stylized decision tree for considering the issues at stake. A potential book author is considering whether to write a book. The possible returns to our author-to-be turn on any number of outcomes that may be difficult to evaluate at the time when the author's fingers first touch the keyboard. One initial question: how successful will the work be? I have captured this question above—somewhat cheekily I confess—in the initial paths in the decision tree: expect to be ignored (professors) and expect to be famous (normal people). I do mean to suggest that I suspect that potential authors suffer from an optimism bias—they believe that they are writing the Great American Novel—though I am not sure that much in my analysis turns on this.

The next step in the tree addresses whether the work will become an orphan. Works that are “remembered” have a good chain of title and there is never doubt as to who owns the copyright. Works that are “forgotten/lost” are the works that become orphans. You can see quickly how the probabilities link up across the events. If you expect your work to succeed, you do not then expect it to be lost. For potential authors who expect their works to succeed, they then, of course, expect title to be tracked carefully, imagining that they think of future title issues at all.

For this group of authors, the orphan works licensing regime is unlikely to play much, if any, of an ex ante role in inducing the creation of works. Putting the same point differently, if there were an ex ante market for orphan works rights, we should expect those rights to go for very little. Of course, if we are dabbling in the behavioral dark arts, then we should also note that the experiments suggest that once you hand people a piece of property—coffee mugs in the experiments and, here, control over subsequent orphan works—status quo bias kicks in and all of a sudden they want a nice chunk of change for the right. But the status quo bias is very much a here-and-now phenomenon—someone standing in front of you offering cash for something you have in your hand—while the orphan works licensing regime is two steps in the tree down the road. Imagining someone demanding meaningful money for the ex ante orphan works right is quite different than imagining that the down-the-road potential licensing rules matter ex ante for our prospective authors.

There are more nodes in our tree and more probabilities to assess. The orphan work needs to be found for it to return to economic significance. It is not fully clear what “found” means. The mass digitization projects will scoop up many orphans but most of them, presumably, will continue to be ignored. The rare use of these orphans might be a technical copyright violation though one could imagine that these uses substitute for uses of the existing physical copies of those works. Those uses—reading and analysis—would, of course, be noninfringing. It is for the orphan works that are meaningfully revived as individual works where we can imagine meaningful disputes over the new economic streams. That takes us to our final event: whether the orphan copyright owner discovers that use.

We can see how distant the orphan works licensing regime is likely to be for our prospective author. You might imagine that not every feature of the copyright law is equally salient at the point that a work was being created. Non-salient features would then be, by definition really, not pivotal or marginal in the creation of works. A designer of a copyright regime would then have free reign in the setting of that particular feature of the copyright rules.

What does that mean for the prospective author? At the time that an author puts fingers to keyboard, how salient should we think the orphan works rules are? Consider the free-use rule and note the attenuated chain of reasoning that seems to matter here. “I won’t write this book now because when my successor copyright holders discover that a book once lost to them is at that point being used by others my successors won’t have a remedy against those users.” To put it mildly, that seems like a big stretch. To some

extent, copyright holders can control whether an orphan works problem arises in the first place through careful management of the copyright.

Presumably, works that turn out to be more valuable are more carefully tracked than less successful works. So as noted earlier, the class of works that we are imagining are works that are insufficiently successful to be worth tracking. Indeed, it is the works that authors imagine at the time of writing to be not worth tracking. If authors have biases that cause them to overestimate the likely success of the work, they will not be concerned about the orphan works rules at all. These authors will expect their books to be successful and will therefore anticipate that they and their successors will manage the work carefully to avoid orphan status.

We could imagine that prospective authors see an option value from their consent right. In that framework, the broader the set of rights given to the author, the more they expect to be able to cash in on undefined and unexpected uses. Those opportunities would be the uses beyond the limits of salience at the time of creating the original work. To take these rights seriously, we would need to get some sense of how the option value of the un-salient is formed and how it changes as we tinker with the scope of exempt uses. One would expect this sort of option value to be quite crude and to be mostly unchanged by changes in the scope of these non-salient rights.

All of that suggests that basing the royalty on the price that is being paid to non-orphan works or that would have been paid in a hypothetical negotiation between the entrant and the copyright holder almost certainly results in a royalty that is too high, as measured by what we want socially. We should expect royalty rates for orphan use to be modest.

V. CONCLUSION

This is a fascinating time in libraries and we are at a key point of institutional design as we transition to digital libraries. We should not want to leave the orphan works outside of this process. At least after the bold failure of the Google Book Search settlement to bring orphan works into wide use,⁵² we seem to need federal legislation to enable full use of the orphans, though we can imagine that fair use may take us part of the way forward.

If we move forward on legislation, we should be sure to take into account recent developments in competition among libraries. We are seeing

52. The settlement was rejected by Judge Chin. *See* http://scholar.google.com/scholar_case?case=11449470169883565960&q=chin+google+book+search+settlement&hl=en&as_sdt=2,14. The merits are now moving to trial.

private digital entrants. These entrants bring with them different business models and they are in many ways more universal and democratic institutions than the world of physical public and nonprofit libraries that have come before them. Online public libraries may face challenges, especially in acquiring new works, and so we should be sure to hedge our bets and make sure that both private and public libraries gain access to the orphan works.

There will be many issues that will arise in creating such legislation and I have focused here on some of the royalty rate issues. The orphan works licensing regime is likely to play very little role in incentivizing the creation of works upfront and that suggests that royalty rates should be low. We need to ensure that orphan works are returned to meaningful use and that may require the creation of a secondary rights regime for the orphans.

THE ORPHANS, THE MARKET, AND THE COPYRIGHT DOGMA: A MODEST SOLUTION FOR A GRAND PROBLEM

Ariel Katz[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1286
II.	PERMISSION-FIRST: THE RULE	1288
III.	PERMISSION-FIRST: THE DOGMA	1290
IV.	COPYRIGHT LAW IS NOT THE COPYRIGHT DOGMA	1293
	A. SEARCH ENGINES AND TECHNOLOGICAL OPT-OUT	1294
V.	REMEDY TWEAKING	1297
VI.	JUSTIFYING REMEDY TWEAKING	1303
	A. WHY DO WORKS BECOME ORPHANS?.....	1304
	B. LEAST-COST AVOIDERS AND LEAST-COST ENABLERS	1306
	C. STICK OR CARROT?.....	1308
	D. “SIT BACK, DO NOTHING” IS RARELY THE LAW.....	1309
	E. JUDICIAL AUTHORITY TO TWEAK REMEDIES.....	1314
VII.	QUASI-DOGMATIC AND DOGMATIC SOLUTIONS	1317
	A. THE U.S. COPYRIGHT OFFICE REMEDY TWEAKING PROPOSAL.....	1318
	B. CANADA’S ORPHAN WORKS REGIME.....	1322
	C. EXTENDED COLLECTIVE LICENSING	1331
	D. DOES ECL EVEN SOLVE THE ORPHAN PROBLEM?	1335
	E. ECL AS INSURANCE.....	1336

© 2012 Ariel Katz.

† Associate Professor, Innovation Chair Electronic Commerce, Faculty of Law, University of Toronto. I wish to thank participants of the Berkeley Orphan Works Symposium (“Berkeley Symposium”), the 2012 Intellectual Property Scholars Conference, and the other events where earlier versions of this Article were presented, for their helpful feedback. I am also grateful to Tun-Jen Chiang, Robert Darnton, Natalie Zemon Davis, Abraham Drassinower, Ed Iacobucci, Lateef Mtima, Simon Stern, and Hal Varian for their insightful comments and suggestions, and to Pamela Samuelson for inviting me to participate in this symposium and her general support and encouragement.

VIII. THE POLITICAL ECONOMY OF THE COPYRIGHT DOGMA	1337
IX. CONCLUSION	1344

I. INTRODUCTION

How do you solve a problem like orphan works? The ability to digitize and make accessible through the internet works that are not commercially available has sparked the imagination of many. At the same time it has also directed the public's attention to the problem of access to and use of works that are still under copyright, but whose copyright owners are unlocatable. One of the most ambitious digitization projects, the Google Book Project, already generated a great deal of controversy and class action litigation, while a proposed settlement to that lawsuit turned out to be even more controversial and was ultimately rejected.¹ The proposed settlement² devised a solution to the orphan works problem, but raised antitrust, privacy, and international law concerns³ and implied, according to many, fundamental changes to copyright law's structure.⁴ With the rejection of the proposed settlement—and with Judge Chin's holding that only Congress is empowered to make such changes⁵—many believe that a legislative change is necessary and that the orphan works problem is one which current copyright law, as implemented by courts, cannot solve.⁶ The thought is that the orphan works problem is grand, and, therefore, requires an equally grand solution.

In this Article, I warn against the quest for finding grand solutions. I analyze some of the proposed solutions and explain why most of them may do more harm than good. I also propose a common law solution to the orphan works problem, based on well-established principles of imposing and limiting liability from other areas of law. I show that courts can implement this solution, in large part, even if Congress takes no action.

Specifically, I propose to address the problem at the remedy stage, rather than by redefining liability. I propose that owners' locatability should be a

1. *Authors Guild v. Google, Inc.*, 770 F. Supp. 2d 666 (S.D.N.Y. 2011).

2. See *Amended Settlement Agreement (ASA)*, GOOGLE BOOK SETTLEMENT, available at <http://www.googlebooksettlement.com/agreement.html> (last visited Aug. 1, 2012).

3. See *Authors Guild*, 770 F. Supp. 2d at 682–86.

4. See, e.g., Pamela Samuelson, *The Google Book Settlement as Copyright Reform*, 2011 WISC. L. REV. 479 (2011).

5. *Authors Guild*, 770 F. Supp. 2d at 677–78 (stating that the issue of orphan works “is better left for Congress”).

6. Pamela Samuelson, *Legislative Alternatives to the Google Book Settlement*, 34 COLUM. J.L. & ARTS 697, 698–700 (2011) (discussing the elements of a “legislative package” that could recreate the goals of the proposed settlement).

mitigating factor that courts should consider when determining the proper remedy for the infringement of a copyright. In appropriate cases courts should deny injunctions and should also have some discretion to deny or reduce statutory damages. The court may still order the user to pay damages, or account for the profit attributable to the infringed work, but if these approximate what the parties would have agreed had a license been negotiated in advance, the orphan works problem would be seriously diminished. While many proposed solutions focus on mechanisms that exempt users of orphan works from liability (and thereby define circumstances and conditions that trigger an immunity),⁷ the proposed solution in this Article (which can be called “remedy tweaking” for convenience) focuses not only on the users’ conduct but also on the owners’ contribution to the problem. Therefore, this solution does not require any change in the basic rules that define when an infringement occurs but solves the problem by tailoring the ensuing remedy.

Remedy tweaking will shift some of the risk of using orphan works from the user to the owner. It will adjust the user’s duty to seek ownership information and permission relative to a corresponding duty of the owner to provide such information. This should motivate copyright owners to make themselves locatable: to de-orphan their works. After all, the orphan works problem is a problem of those copyright owners’ own making, and they are the ones who can mitigate it at the lowest cost. Tort law takes into account a plaintiff’s contributory or comparative negligence; contract law demands that plaintiff take reasonable effort to mitigate their damage before they can recover its full extent. Property law sometimes imposes a duty to provide notice before a right can be enforced. The law does not recognize an unfettered right to sit back and do nothing in all circumstances and regardless of the consequences. Copyright law should not recognize such a right either.

I also discuss and warn against solutions that promote monopolies, such as those involving collective licensing or “extended collective licensing.” The underlying assumption of this Article is that, at its core, copyright law is based on a very simple logic—market logic. Copyright law envisions a decentralized market system, whereby the law grants limited exclusive rights in creative works, with the expectation that such rights will then be voluntarily exchanged in a decentralized market place.

The idea of a decentralized marketplace lies at the heart of copyright law. Indeed, this idea was encoded in the DNA of copyright since its statutory

7. See discussion *infra* Part VII.

inception. The *Statute of Anne*, the first copyright statute, arose not only out of the publishers' demands for statutorily protected exclusive rights, but also reflected the English Parliament's disdain of the monopoly that the Stationers' Company—the self-regulating cartel of London's publishers—had over the book trade. For nearly two centuries following arrival of the printing press in Britain, the London Stationers, alongside the Crown and the Cross, attempted to control the press and the book trade.⁸ The combination of monopoly and press control was an early reaction to the printing press by those who felt threatened by the ways it revolutionized how knowledge was being produced, disseminated, and controlled.⁹ At the turn of the eighteenth century, however, Parliament rejected such control, and modern copyright law was born. We should keep this point in mind when we design the rules governing the dissemination of knowledge as society moves from print to digital. And when facing specific problems, we should be cautious not to adopt solutions that could bring back a new version of the Stationers' Company and leave monopolies and bureaucracies as gatekeepers of orphan works and knowledge.

Part II explores the Permission-First Rule as the default rule in copyright law. Part III further discusses how the Permission-First Rule has become a dogma for copyright purists, and Part IV examines traditional deviations from the Permission-First dogma. Part V introduces “remedy tweaking” as a possible solution to the orphan works problem. Part VI continues with a discussion of how “remedy tweaking” is not only useful, but also justified. Part VII examines and critiques current “dogmatic” solutions to the orphan works problem. Part VIII discusses the political economy of the copyright dogma: who stands to benefit from promoting it, how the public stands to lose, and how those interests influence the proposed solutions. Part IX offers a brief conclusion.

II. PERMISSION-FIRST: THE RULE

As a general rule in copyright, a user¹⁰ should contact the owner and seek permission *before* using the work. Let us call it the permission-first rule.¹¹

8. See, e.g., JOSEPH LOEWENSTEIN, *THE AUTHOR'S DUE: PRINTING AND THE PREHISTORY OF COPYRIGHT* 14 (2002); LYMAN RAY PATTERSON, *COPYRIGHT IN HISTORICAL PERSPECTIVE* 152–91 (1968).

9. See PATTERSON, *supra* note 8, at 152–91.

10. I use terms “user” and “use” for convenience, even though copyright law does not grant exclusive right to *use* a work, only exclusive rights with respect to specific uses such as reproduction and public performance. See 17 U.S.C. § 106 (2010).

11. Of course, “permission-first rule” is just a synonym for the now-standard term “property rule” (as distinct from a “liability rule”) as used in the law and economics

Generally in copyright, permission-first is a sensible rule—ordinarily it is much less costly for prospective users to ask permission from the owner, than for an owner to locate all actual or potential users and ask them to pay.¹² In the paradigmatic case, the owner is singular, locatable, and stable, whereas users could be many and varied, scattered and fleeting.¹³ Moreover, requiring the putative user to approach the owner and seek permission forces the user to disclose the ways in which he plans to use the work; this, in turn, makes it easier for the owner to tailor the appropriate license fees and conditions through various forms of price discrimination.¹⁴ Therefore, a permission-first rule (and strict liability for violating it) generally supports copyright's market orientation.

But generally is not always, and not every case is paradigmatic.¹⁵ A permission-first rule only makes sense if the cost of getting prior permission is low, or at least low relative to the expected value of using the work, and if it is less costly for the user to seek prior permission than for the owner to

literature, following Calabresi & Melamed. See Guido Calabresi & Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972). However, since the term “property” could have a broader meaning in law, outside law and economics, I prefer using “permission-first” in this paper to emphasize this particular aspect of the copyright entitlement.

12. See, e.g., Oren Bracha, *Standing Copyright Law on Its Head – The Googlization of Everything and the Many Faces of Property*, 85 TEX. L. REV. 1799, 1827 (2007); Antonio Nicita & Giovanni B. Ramello, *Property, Liability and Market Power: The Antitrust Side of Copyright*, 3 REV. L. & ECON. 767, 775 (2007).

13. The non-rivalrous nature of creative works exacerbates the problem because any user who has access to a work can practically use it without negotiating with the owner at all. Under a permission-later rule, some users would use without ever getting permission, hoping that their use would go unnoticed, and using any stalling tactic to delay payment once contacted.

14. See Wendy J. Gordon, *Intellectual Property As Price Discrimination: Implications For Contract*, 73 CHI-KENT L. REV. 1367, 1371–72 (1998).

15. While strict liability for copyright infringement has garnered wide support by courts and commentators, some commentators have offered more nuanced views and argued against its uniform applicability. See, e.g., Dane S. Ciolino & Erin A. Donelon, *Questioning Strict Liability in Copyright*, 54 RUTGERS L. REV. 351 (2002); Avihay Dorfman & Assaf Jacob, *Copyright as Tort*, 12 THEORETICAL INQ. L. 59, 77–82 (2011) (arguing that “the same considerations that call for the protection of tangible property through a regime of strict liability torts require (or may often require) abandoning this regime in the case of copyright” and replacing it with rules similar to those that apply to accidents, “which may include the owners’ contributory or comparative negligence”); see also T. J. Chiang, *The Reciprocity of Search*, 66 VAND. L. REV. 1 (2013) (critiquing the way patent law, like copyright law, imposes the duty to locate patent owners solely on users). Chiang argues that, contrary to conventional wisdom, patentees are often better positioned to locate users and that the law should impose a duty to search on patentees in such cases. *Id.* Chiang also mentions the orphan work problem as a parallel problem in copyright law. *Id.* at 60–63.

contact the user.¹⁶ If the cost of obtaining prior permission becomes prohibitively high relative to the contemplated use, then a system demanding prior permission would result in a market failure. The problem with many mass digitization projects—especially when they involve orphan, partly orphan, or quasi-orphan works¹⁷—is that seeking permission first is either impossible or prohibitively costly.¹⁸ Therefore, in such cases permission-first is not a workable rule. In such cases we need an alternative rule.¹⁹

III. PERMISSION-FIRST: THE DOGMA

Unfortunately, solving the orphan works problem has been much more difficult than acknowledging its existence. One of the main obstacles has been a tendency to treat the permission-first rule as a dogma—“a belief held unquestioningly and with undefended certainty”²⁰—instead of a rule whose validity should always be judged in light of the purpose of copyright law. For some copyright purists, the rule should be adhered to irrespective of circumstances and regardless of the consequences because “permission-first” is the essence of copyright, and anything else would turn copyright “on its

16. See Chiang, *supra* note 15.

17. A completely orphan work would be one where the owner is unknown and unlocatable. A work can be partly orphan if the owner is known but unlocatable (e.g., it is known that the author is John Smith and that he is American, but there are 45,352 Americans named John Smith, see HOWMANYOFME, <http://www.HowManyOfMe.com> (last visited May 21, 2012)), or locatable but unknown (e.g., the work was written by an anonymous writer or under a pseudonym, although it would not be difficult to locate the author if only one knew her identity). By quasi-orphan work I mean a work in which ownership has changed without sufficient record, or a work with fragmented ownership which, practically, makes it prohibitively costly to negotiate a license.

18. The problem is by no means limited to digitization. As Eric Schwartz explained in his presentation at this symposium, the orphan state of many films is a challenge for efforts to preserve them, and presently, most preservation is done by non-digital means. Eric Schwartz, Panel Discussion, *Orphan Works: Films and Recordings* at the Berkeley Symposium: “Orphan Works and Mass Digitalization: Obstacles and Opportunities” (Apr. 12, 2012), available at <http://www.law.berkeley.edu/11731.htm>. Moreover, the orphan works problem is barely a new one. Indeed, the Ottoman Empire’s copyright legislation from 1910 included a special rule applied to orphan works whose authors died and who had no heirs—the work could be copied, published and translated. See MICHAEL D. BIRNHACK, COLONIAL COPYRIGHT: INTELLECTUAL PROPERTY IN MANDATE PALESTINE 85 (2012).

19. Fair use, which exempts the user from obtaining any permission, is sometimes explained as a rule designed to prevent such market failure. See, e.g., *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 592 (1994) (reasoning that the “unlikelihood that creators of imaginative works will license critical reviews or lampoons of their own productions” justifies treating parodies as fair use).

20. SIMON BLACKBURN, THE OXFORD DICTIONARY OF PHILOSOPHY 104 (2d ed. 2005).

head.”²¹ This view of copyright appears to have influenced Judge Chin in his decision to reject the proposed settlement in the Google case.²² Judge Chin endorsed a statement by Amazon’s counsel, David Nimmer, that “[t]he law of the United States is a copyright owner may sit back, do nothing and enjoy his property rights untrammelled by others exploiting his works without permission.”²³ Sit back, do nothing, regardless of the consequences. The dogma.

For the copyright dogma, the wrong in using someone else’s work without permission has nothing to do with impairing the ability of the copyright owner to participate in the marketplace or with depriving the owner of the opportunity to share the benefits resulting from the use of her work. Nor does it lie in any other measurable concept of harm to the owner or to society. If it did, there would be no problem recognizing that the permission-first rule could be relaxed when the owner is unlocatable or when the use cannot cause any harm and the rule prevents beneficial uses of the work.

The copyright dogma does not treat copying without permission as a harmful act against an owner but, rather, as a sinful act, wrongful in and of itself. The dogma does not regard the use of works as something that ought to be celebrated. It does not seek to justify copyright’s limitations on use by the need to reward creators and encourage creativity, or to recognize authorial dignity. The dogma denies “that copyright law is less a prohibition on copying, or an obligation not to copy, than an institutionalized distinction

21. See Bracha, *supra* note 12, at 1802 (explaining the ideology of Google’s critics that copyright owners should not “carry the burden” of enforcing their rights).

22. *Authors Guild v. Google, Inc.*, 770 F. Supp. 2d 666, 669–70 (S.D.N.Y. 2011).

23. *Id.* at 681. Judge Chin viewed this conclusion as a natural corollary to the undisputed proposition that “[a] copyright owner’s right to exclude others from using his property is fundamental.” *Id.* (citing *Fox Film Corp. v. Doyal*, 286 U.S. 123, 127 (1932)). But he did not provide any reference to the proposition that this fundamental right to exclude bears no exceptions whatsoever. In fact, if, as the Court held in *Fox Film*, the “primary object in conferring the monopoly lie[s] in the general benefits derived by the public from the labors of authors,” then arguably, the ability to sit back and do nothing may be limited if it deprives the public from the very benefits that the copyright was designed to foster. See *Fox Film*, 286 U.S. at 127. In fairness to Judge Chin, however, the question he confronted was not whether there could be any circumstances justifying deviation from the permission-first rule, but whether to approve a highly controversial settlement that would involve a massive reshuffling of rights and the setting up of new institutions—a task which he thought would be better left for Congress. *Authors Guild*, 770 F. Supp. 2d at 677. Moreover, Judge Chin was not asked to consider (or reconsider) the dogma but, rather, to choose among competing interpretations of it. The lawsuit against Google, the Amended Settlement Agreement (ASA), and many of the objections to the ASA, were all reflections of the dogma. And like them, many proposed solutions to the orphan works problem are captive to it.

between permissible and impermissible copying.”²⁴ Therefore, in attempting to determine when the law should require prior permission the dogma does not use a matrix of considerations aiming to optimize reward, incentives, dissemination, freedom, dignity, or any such or other factors. The copyright dogma does not define copyright by the rights it confers on its owner and the correlative duties it imposes on others; it does not ask how those correlative rights and duties can be exercised and complied with in ways that promote the public interest, minimize interference with the rights and liberties and others, or otherwise achieve copyright law’s purpose.²⁵ Instead, it defines copyright solely by the prohibitions it imposes on users. Thus, the dogma views unauthorized copying as inherently sinful, and regards prior permission as the means to avoid and absolve that venial sin of copying.

When users and owners can easily obtain and grant permissions, debating the nature of the wrong in copyright may seem interesting but of little consequence. However, the question becomes crucial when transactions are prohibitively costly. Such situations pose an immediate challenge to the dogma. Obviously, if the copyright owner is unlocatable, there would be no mechanism to absolve the sin of copying. Therefore, to avoid sin, the dogma must insist that the work must never be copied. While this position may be internally consistent, it would not garner much political support if it sentences orphan works to oblivion. Consequently, to retain some of its appeal, the dogma must provide substitutions to the absent owner: someone else who can give permission (and collect a fee) in lieu of the absent owner.

The doctrinal move is less challenging than it initially appears. In fact, once the sin is detached from any concept of harm to the copyright owner and becomes the lack of permission itself, it no longer matters whether permission is sought from the owner or from someone else, as long as that person is willing to step into the shoes of the absent owner and collect a fee in her stead. In fact, one cannot help but suspect that those who could benefit from stepping into the shoes of the absent owners are those who have the greatest motivation to promote the dogma.²⁶

This kind of development of the copyright dogma and the solutions that it proposes resemble some aspects of Catholic dogma in the Middle Ages; the Catholic Church realized that if it could not prevent sin altogether, it

24. See Abraham Drassinower, *Copyright is Not About Copying*, 125 HARV. L. REV. 108, 112 (2012).

25. See *id.* at 109–10; see also Shyamkrishna Balganesh, *The Obligatory Structure of Copyright Law: Unbundling the Wrong of Copying*, 125 HARV. L. REV. 1664 (2012).

26. See *infra* Part VIII.

might as well monetize sin through the sale of indulgences.²⁷ The similarities between the dogma's solutions and indulgences will be discussed later in the Article, *infra*. At this point, however, my purpose is to show that copyright law, unlike copyright dogma, is not very dogmatic at all. Indeed, deviations from the permission-first rule are part and parcel of copyright law.

IV. COPYRIGHT LAW IS NOT THE COPYRIGHT DOGMA

Consider fair use and other exceptions to copyright. Fair use allows users, with some limitations, to copy protected works for legitimate purposes.²⁸ Therefore, fair use “permits and requires courts to avoid rigid application of the copyright statute when, on occasion, it would stifle the very creativity which that law is designed to foster.”²⁹ Creating an exception has sometimes been viewed as a solution that copyright law applies to situations in which requiring owners to get permission would result in a market failure.³⁰ For example, it is very hard to imagine a functioning market for rights to criticize the works of others.³¹ This Article does not discuss fair use here to argue that fair use *is* the solution for the orphan works problem (although in some cases it certainly might be).³² Rather, the Article only mentions fair use to demonstrate that copyright law has departed from the permission-first rule when necessary. In fact, fair use goes beyond merely deviating from the permission-first rule because, when it applies, the user does not need to seek permission at all.³³

27. *Id.*

28. 17 U.S.C. § 107 (2010) (“[T]he fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright.”).

29. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 577 (1994) (internal quotation marks and citation omitted).

30. See Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and its Predecessors*, 82 COLUM. L. REV. 1600, 1604–05 (1982).

31. See WILLIAM M. LANDES & RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW* 153 (2003); *Campbell*, 510 U.S. at 592 (explaining the distinction between “biting criticism that merely suppresses demand and copyright infringement, which usurps it”) (internal quotations and punctuation omitted).

32. See Jennifer M. Urban, *How Fair Use Can Help Solve the Orphan Works Problem*, 27 BERKELEY TECH. L. J. 1379 (2012); see also William F. Patry & Richard A. Posner, *Fair Use and Statutory Reform in the Wake of Eldred*, 92 CALIF. L. REV. 1639, 1639–61 (2004). The recent ruling in *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939 (S.D.N.Y. Oct. 10, 2012), may indeed indicate that fair use can be relied on to facilitate many digitization projects involving all works, including orphan works.

33. In fact, as others have noted, this feature of fair use is sometimes an issue because it can lead to extreme outcomes: strict liability if a defendant's fair use is rejected or zero

There are, however, situations in which a duty to seek permission first (ex ante) might lead to a market failure—or a highly inefficient market—even though there might be room for negotiation ex post. In such cases, the law may not necessarily provide total exemption in the way that fair use does,³⁴ but it may change the default rule. The law may shield a user from liability for using a work without getting prior permission until the copyright owner objects. The notice-and-take-down mechanism for online service providers (“OSPs”) under § 512(c) of the Copyright Act³⁵ is an example of such a mechanism. It creates certain limitations on the liability of OSPs for copyright infringement of their subscribers. The mechanism creates a safe harbor that allows OSPs to host the infringing material until the copyright owner provides notice of the unauthorized use.³⁶ The notice-and-take-down mechanism does not endorse a “sit back, do nothing” approach. In fact, it conditions some of the copyright owners’ entitlements on taking specific actions in prescribed form.³⁷ Section 512(c) is but one example in which the law deviates from the “permission-first” rule. This particular rule immunizes OSPs from potential contributory or vicarious liability for the direct infringement of their subscribers. The next Section discusses an example of deviation from the “permission-first” rule that applies to direct infringers.

A. SEARCH ENGINES AND TECHNOLOGICAL OPT-OUT

The growing judicial endorsement of technological opt-out solutions, such as those applied by search engines, reflects another instance of courts’ willingness to adopt solutions that deviate from a strict permission-first rule, but do not go as far as complete exemption.³⁸

Consider the copyright issues that Google and other search engines face. In order to run its search engine, Google systematically copies every webpage on the Internet, stores it on its servers, and indexes it. And Google copies

liability if successful, without anything in between. See Alex Kozinski & Christopher M. Newman, *What’s So Fair About Fair Use? The 1999 Donald C. Brace Memorial Lecture*, 46 J. COPYRIGHT SOC’Y U.S.A. 513, 515 (1999) (“In fair use, as with pregnancy tests, ‘a little bit’ isn’t considered an acceptable response. Fair use is conceptually a hard-edged box; either you’re in it or you’re out of it. If you’re out, you can be enjoined out of existence.”).

34. See 17 U.S.C. § 107 (2010) (“[T]he fair use of a copyrighted work . . . is not an infringement of copyright.”).

35. 17 U.S.C. § 512(c) (2010).

36. See *id.*

37. See Peter Dicola & Matthew Sag, *An Information-Gathering Approach to Copyright Policy*, 34 CARDOZO L. REV. 173, 183 (2012) (explaining the pervasiveness of “tolerated use” in social media contexts because copyright owners must provide a notice of alleged infringement).

38. See, e.g., *Field v. Google Inc.*, 412 F. Supp. 2d 1106 (D. Nev. 2006); *Parker v. Yahoo!, Inc.*, No. 07-2757, 2008 WL 4410095 (E.D. Pa. 2008).

every such webpage without asking permission first. Obviously, Google could not personally contact every website owner and get her permission before indexing her website in a manner that was not extremely costly and inefficient.³⁹

Instead, Google believes that it can use an opt-out system. Google employs “Googlebot,” a computer program, which, by default, copies every webpage and extracts from it the necessary data.⁴⁰ Google does not seek permission first, but it does not copy and index a webpage against the website owner’s wishes.⁴¹ To prevent unauthorized use, Google (and other search engines) rely on a technological opt-out system. This opt-out system allows website owners to automatically notify Google if they do not want their website or specific parts of it to be indexed.⁴² Website developers can simply insert widely available lines into the html code of the website (known as “Robots exclusion protocol” or robots.txt).⁴³ Google’s computers that otherwise would automatically copy everything recognize the code and know to keep out.⁴⁴ Therefore, Google’s business model relies on the theory that since strict compliance with a permission-first rule would be extremely inefficient, and since website owners are able to communicate their desires in a very simple, cheap, and effective way, it does not need to get permission first.

The legality of Google’s theory has not been fully tested or universally accepted, and has on occasion been contested.⁴⁵ With some exceptions,⁴⁶

39. See Bracha, *supra* note 12, at 1827. Note, however, that it is not strictly impossible for Google to do that. In theory, Google could employ some people to visit websites, find their owner’s contact details, and contact them to seek their permission to index them, but obviously this would be extremely inefficient.

40. *Googlebot Webmaster Tools Help*, GOOGLE, <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=182072> (last visited Apr. 20, 2012).

41. See Bracha, *supra* note 12, at 1827.

42. *Id.*

43. See About /robots.txt, ROBOTSTXT.ORG, <http://www.robotstxt.org/robotstxt.html> (explaining the robots exclusion protocol) (last visited Nov. 30, 2012).

44. *Id.*

45. See *Google, Inc. v. Copiepresse, Cours d’Appel [CA]* (Court of Appeal) Bruxelles, 9e ch. May 9, 2011, (Belg.), available at http://www.copiepresse.be/pdf/Copiepresse%20-%20ruling%20appel%20Google_5May2011.pdf (rejecting Google’s argument that a failure to implement robots.txt protocols constituted permission to copy and index newspaper). The court reasoned that “it cannot be permitted that a holder is deprived of his rights simply because he has neglected to implement a technological process or, as [one of the respondents] so colorfully puts it, that it would be ‘legal to rob a house of its contents because a door was left open!’” *Id.* ¶ 50.

46. See *id.*

courts have tended to be sympathetic towards this theory.⁴⁷ However, it is important to note that this solution departs from the permission-first rule. It requires website owners to take concrete steps to signal the lack of permission. Essentially, this solution not only allows Google to say “we will copy you, unless you tell us not to,” but it also allows Google to add “and unless you tell us not to in a very specific way.” If the theory is upheld—and considering how rejecting it would thwart the utility of search engines one can hope that it will—it can teach an important lesson: when a permission-first rule is highly inefficient, an alternative system that imposes a minimal burden on copyright owners, while remaining responsive to and respectful of their wishes, might be acceptable. The jurisdictions that endorse Google’s practice for indexing websites clearly reject the notion that website owners, when they have copyright in the content of their website, can “sit back, do nothing” and have their website untouched. If they do not want Google to index their websites, they need to take action, and until they do, they will not have recourse against Google.⁴⁸

Obviously, the technological opt-out solution that works for internet websites is not directly applicable to books that were not born digital; it also may not be available to many other mass-digitization projects and other uses of orphan and non-orphan works. Unlike websites, the “code” of printed books cannot contain robots.txt. But a general principle can be distilled from the technological opt-out mechanism that can apply more broadly, from forgotten orphans to living websites: when a permission-first rule leads to a market failure, or results in marked inefficiency, the law may deviate from permission-first. Further, if there is an alternative to the permission-first rule that, by imposing a minimal burden on owners, while remaining responsive to and respectful of their wishes, prevents the market failure or avoids such inefficiency, the law should endorse such a rule.

47. See, e.g., *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1109 (D. Nev. 2006) (granting Google’s motion for summary judgment); see also Arnd Haller, *German Supreme Court: Google Image Search in Line with Copyright Law*, EUR. PUB. POLY BLOG (Apr. 11, 2012, 10:53 AM), <http://googlepolicyeuropa.blogspot.com/2012/04/german-supreme-court-google-image.html> (discussing a decision by the German Supreme Court, which upheld the legality of Google’s indexing methodology).

48. The implied assumption here is that Google has to be responsive to copyright owners’ wishes when it copies their websites for the purpose of indexing them. There is, however, a credible argument that this kind of copying for non-expressive purposes does not infringe copyright at all. See Matthew Sag, *Orphan Works as Grist for the Data Mill*, 27 BERKELEY TECH. L.J. 1503 (2012). The implication of this view is that as a matter of copyright law, copying websites in order to index their content is a form of non-expressive use that ought to be allowed even against the wishes of copyright owners.

In the context of search engines, this principle manifests itself in the form of a technological opt-out solution. But legally, and abstracting from the technological particulars of the solution, the principle can translate into a more general recognition that while users have a duty to seek permission before using—and not to engage in protected uses of the work against the owner’s wishes—owners may have a corresponding duty to take reasonable steps to facilitate users’ ability to seek permission or otherwise communicate their wishes efficiently. While failure to fulfill the user’s duty would constitute an infringement of the copyright, failure to comply with the owner’s duty might deprive her from some of the remedies to which she otherwise would be entitled.

V. REMEDY TWEAKING

The previous Part discussed some contexts where the copyright owner’s right or her entitlement to the full range of remedies might be conditioned on communicating her wishes to users effectively. From that discussion emerged a broader and more general principle: while a user has a duty to refrain from engaging in protected uses of a work against its owner’s wishes, the owner has a corresponding duty to take some steps to facilitate the user’s ability to seek permission or otherwise communicate their wishes efficiently. A user who breaches his duty is an infringer, but an owner who fails to fulfill her part may be barred from certain types of relief. This Part applies the principle of owner responsibility to orphan works, and shows that the principle, as applied, provides a satisfactory solution to the problem of orphan works. Moreover, the discussion in this Part indicates that for the most part, this solution is within grasp. It requires neither “Grand Solutions,” as followers of the dogma propose, nor substantial legislative reform. I call this solution “remedy tweaking.”

The proposal is simple. The orphan works problem can be addressed by tweaking the remedies available for infringement. The locatability of the owner—the ease or difficulty of locating the owner—should be a factor that the court would consider when exercising its discretion in tailoring an appropriate remedy. This solution is similar in some respects to the proposal made by the Copyright Office in 2006⁴⁹ but different in others.⁵⁰ Rather than changing what constitutes an infringement, the problem of orphan works can be addressed by modifying the consequences arising from infringement.

49. REGISTER OF COPYRIGHTS, REPORT ON ORPHAN WORKS (2006), *available at* <http://www.copyright.gov/orphan/orphan-report-full.pdf>.

50. *See infra* Section VII.A.

Remedy tweaking will shift some of the risk of using orphan works from the user to the owner. It will adjust the user's duty to seek ownership information and permission relative to the owner's corresponding duty to provide such information. This solution will achieve two goals. First, it will remove some impediments to using the remaining orphan works because it will reduce the risk of disproportionate penalties. Second, it may shrink the number of orphan works by motivating copyright owners to make themselves locatable and "de-orphan" their works, and making it easier for users to seek permissions and obtain them.

The proposed solution does not change any of the basic rules of liability. Therefore, anyone who violates the owner's exclusive right is an infringer,⁵¹ regardless of whether the owner was locatable. Instead, the proposed solution takes into account the copier's ability to locate the owner and seek prior permission when the court considers the appropriate relief that the owner will be granted. When a defendant uses the work without permission and under circumstances where consent could have been easily obtained, the copyright owner should be entitled to the full range of remedies. However, when permission could not have been easily sought, the court may adjust the remedy. Thus, a court's inquiry should not be limited to asking "did the user seek permission before copying?", but should also include questions such as: "did the user make a good faith attempt to seek permission?", "what steps were taken?", and "what steps could have been taken but were not?". However, even if the court finds that the defendant could have taken additional measures that would have resulted in locating the owner, the court should also consider whether those measure *should* have been taken and whether they should have been taken prior to using the work or, perhaps, after.⁵² In other words, what kinds of measures would be considered "reasonable" in the particular case? In this regard, the proposal is not limited to orphan works or to situations where the owner is unlocatable, but can apply to a wider range of situations, such as when locating the owner in advance is technically possible, but extremely inefficient.⁵³

51. See 17 U.S.C. § 501(a) (2010).

52. Stef Van Gompel, Panel Discussion, *The Orphan Works Chimera and How to Defeat It: A View from Across the Atlantic* at the Berkeley Symposium: "Orphan Works and Mass Digitalization: Obstacles and Opportunities" (Apr. 12, 2012), available at <http://www.law.berkeley.edu/11731.htm>. In his presentation, Stef Van Gompel described how TV producers in the Netherlands sometimes add a notice at the end of the program announcing that they have tried to clear all necessary copyrights and invite anyone who believes that his work was used without permission to contact the producer.

53. As in the case of search engines, see *supra* Section IV.A.

The answers to those questions may depend on different factors, such as the characteristics of the user, the nature of his project, the kind of work he copied, the availability of effective search tools and technologies, and so on. What may be reasonable in one case may not be reasonable in another. In addition, the court's inquiry should not only focus on the *user's* efforts to seek permission but also on the *owner's* efforts: what measures he took, could have taken, and *should* have taken to make it easy for prospective users to seek permission. The answers to these questions should also depend on various factors, such as the characteristics of the owner, the type of work, the tools and technologies that were available for recording and communicating ownership information to potential or actual users, and so on.

For example, a Hollywood studio producing a commercial film based on an orphan novel should be expected to exert more effort in locating the owner than a film student producing a film based on the novel. Or a producer of a commercial feature film should be expected to exert more effort than a producer of a low-budget documentary. If the name of the author or the publisher appear on the novel, the copyright was registered, and consulting the registry would easily lead to the current owner, then the differences between the productions and the producers might have little effect on the owner's entitlement to the ordinary range of remedies. But if the work in question is an old reprint of a photograph that plays a minor role in the film and bears no information whatsoever about the identity of the photographer or the copyright owner, then a limitation on the remedy might be equally appropriate whether the user is a Hollywood studio or a student. Yet, if the photo is new, and technologies exist that allow embedding ownership information in the photo itself (or embed an identifier that can be expeditiously matched against an accessible and reliable ownership database),⁵⁴ then failure to contact the owner should weigh heavily against the user, while failure to embed this information should weigh against the owner.

Another consideration affecting the choice of remedy should be the timing of the lawsuit—whether action was brought before the work has been embedded in another project or after, and whether it can be easily taken out or replaced. Courts should consider these factors in determining whether granting an injunction would give the owner disproportionate leverage that would allow her to demand a settlement payment that far exceeds what would have been negotiated *ex ante*.⁵⁵ Therefore, when an injunction would

54. See Molly Shaffer Van Houweling, *Author Autonomy and Atomism in Copyright Law*, 96 VA. L. REV. 549, 557–58, 631–32 (2010).

55. See Chiang, *supra* note 15, at 7–8 (describing the similar holdup problem in the context of patents).

give the owner a disproportionate holdup power, the court may deny an injunction, and limit the remedy to an award of damages.⁵⁶ That courts are permitted to deny injunctive relief is currently beyond dispute.⁵⁷

The possibility of injunctive relief may be a serious deterrent for some users contemplating the use of an orphan work, especially if the work cannot be taken out or replaced (e.g., a movie based on an orphan novel). But if the court only awards damages approximating what the parties would have agreed to had a license been negotiated in advance, the orphan works problem would be seriously diminished. As the Article discusses in more detail, *infra*, most orphan works are presumably those with very low expected commercial value; they presumably are not the ones that would have allowed the owner to bargain for any significant license fees *ex ante*. For some uses, there would be many close substitutes to any single orphan work, meaning that any license fees negotiated *ex ante* would be close to zero,⁵⁸ although this may not be the case in all types of uses. For example, users engaged in non-fictional projects may wish to incorporate specific works to which no other work may be a substitute, whereas projects such as digital archives, indexes, or libraries may wish to include as many works as possible. The value of such products often increases the more comprehensive they are,

56. In fact, as Hal Varian explains, copyright owners' holdup power may lead users to refrain from seeking permission and using works even when the owners are locatable. Hal R. Varian, *Copyright Term Extension and Orphan Works*, 15 INDUS. & CORP. CHANGE 965, 972 (2006), available at <http://icc.oxfordjournals.org/content/15/6/965.full.pdf>. That is because the "effort exerted by the buyer is a sunk cost; there is no reason for the seller to take it into account in setting a price. But the result is that the buyer can never recover its costs of search, leading it to avoid even trying." *Id.* This is another reason why remedy tweaking may be an approach that should be used more broadly, not only for works that are designated as orphaned.

57. *See* eBay Inc. v. MercExchange, L.L.C., 547 U.S. 388, 392–93 (2006) ("Like the Patent Act, the Copyright Act provides that courts 'may' grant injunctive relief 'on such terms as it may deem reasonable to prevent or restrain infringement of a copyright.' And as in our decision today, this Court has consistently rejected invitations to replace traditional equitable considerations with a rule that an injunction automatically follows a determination that a copyright has been infringed.") (internal citations omitted).

58. *Cf.* Ty Inc. v. Perryman, 306 F.3d 509, 512 (7th Cir. 2002) (explaining the concept of "free riding" in the trademark context). Judge Posner remarks that it is unlikely that the owner of a prestigious trademark could obtain substantial license fees where there is no consumer confusion or brand tarnishment because competition would drive the fee to zero; if the name is being used in an unrelated market, virtually every prestigious name would be a substitute for every other in that market. *Id.* In the copyright context, consider, for example, a movie scene in which a protagonist is sifting through old family photographs—obviously, there could be numerous photographs that could be equally suitable for inclusion, not to mention the ability to digitally produce old-looking photograph with Photoshop or similar computer programs.

meaning that the individual works are complements and not mere substitutes.⁵⁹ In any event, the fact that the owner has let the work become orphaned should be viewed as a concession that she did not expect to be able to command any significant license fees.

An infringement of copyright may normally entitle the owner to receive the profit accrued to the user and attributable to the infringed work.⁶⁰ The risk (and the resulting potential for over-deterrence) that accounting for profit may pose depends on the nature of the user's project. For example, projects such as libraries, archives, or indexes may be highly valuable (and profitable—think Google), but the value (and profit) that can be attributable to any single work may be quite small. Subject to the risk of statutory damages and class actions, which will be addressed below, the risk of accounting for profit may not be a significant deterrent for pursuing these projects. In other types of projects, such as recorded songs, movies, or other derivative works, the profit attributable to the infringed work might be quite high (or viewed as such to a judge or jury). In such cases, using an orphan work might be risky. Limiting the amounts that the copyright owner may recover will reduce the risk.⁶¹

Moreover, although in some cases denying an injunction and opting for a tweaked monetary award may reduce the risk associated with using orphan works, in other cases—especially those involving non-commercial activities—injunctive relief might be more appropriate than a monetary remedy. The purpose of this Article is to propose remedy tweaking in broad brush, not to work out its specific details; the idiosyncrasies of the many users, uses, works, and owners suggest that working out the optimal solution requires a case-by-case analysis better left to the courts.

The crucial point is that remedy tweaking will shift some of the risk of using orphan works from the user to the owner. This, in turn, will motivate copyright owners to make themselves locatable—to de-orphan their works—

59. Whether this complementarity would allow the owner to exercise any market power depends on how much withdrawing the work might decrease the project's total value. Even if no individual work may allow the owner to hold up the project completely, even a small amount of fees per work may accumulate to a significant amount that exceeds the user's ability to pay. This might be a justification for the favorable treatment of libraries and archives under copyright law, or broad fair use rights for such projects, although this exceeds the scope of this Article.

60. See 17 U.S.C. § 504(b) (2010) (“The copyright owner is entitled to recover the actual damages suffered by him or her as a result of the infringement, and any profits of the infringer that are attributable to the infringement.”).

61. Section VI.D, *infra*, discusses whether courts have the power to tweak such remedies.

and encourage users to revive orphan works when their owners are still unlocatable. Since courts have wide discretion to tailor the right remedy for the individual case, they can tweak the remedies when appropriate even if Congress takes no legislative action on the issue of orphan works. Of course, Congress can adopt legislation that encourages courts to exercise this discretion, or simply clarifies that such discretion exists. But with two possible exceptions, remedy tweaking, unlike other proposed solutions, does not depend on statutory reform.

One exception is statutory damages. If courts indeed do not have discretion to deny or reduce them below the statutory minimum, statutory damages may remain a significant problem.⁶² Without such flexibility the remedy may not be fully tweakable. Therefore, statutory reform that ensures judicial discretion would be fruitful.⁶³

The other exception is class actions. Remedy tweaking may work well when users contemplate using one or just a few works (or more accurately, when the use involves one or just a few claims of infringement). In such

62. As others have noted, the award of statutory damages—that can fairly be deemed “grossly excessive”—may violate due process of law. See J. Cam Barker, *Grossly Excessive Penalties in the Battle Against Illegal File-Sharing: The Troubling Effects of Aggregating Minimum Statutory Damages for Copyright Infringement*, 83 TEX. L. REV. 525, 538–40 (2004) (discussing the history of the Supreme Court’s approach to excessive punitive damages); see also Pamela Samuelson & Tara Wheatland, *Statutory Damages in Copyright Law: A Remedy in Need of Reform*, 51 WM. & MARY L. REV. 439, 480–91 (2009) (discussing numerous cases in which the damage awards seem inconsistent with legislative intent).

63. See Barker, *supra* note 62, at 558 (advocating for statutory reform to create more “nuanced apportionment” of damages); Samuelson & Wheatland, *supra* note 62, at 509–10 (arguing for reforms that grant courts the ability to reduce statutory damages in certain contexts and limit the availability of statutory damage awards, overall). Although this point does not seem to have been raised in the American literature on statutory damages in copyright law, it is arguable that, at least in some common law jurisdictions, courts possess an inherent power to grant relief against penalties and forfeitures and deny an award that is grossly disproportionate to the plaintiff’s harm and to the defendant’s blame. For example, Section 98 of the Ontario Courts of Justice Act, provides that “[a] court may grant relief against penalties and forfeitures, on such terms as to compensation or otherwise as are considered just.” Courts of Justice Act, R.S.O. 1990, c. C.43, s. 98 (Can.), available at <http://canlii.ca/t/ldxl>. The Supreme Court of Canada explained that “[t]he power to grant relief against forfeiture is an equitable remedy and is purely discretionary. The factors to be considered by the Court in the exercise of its discretion are the conduct of the applicant, the gravity of the breaches, and the disparity between the value of the property forfeited and the damage caused by the breach.” See *Saskatchewan River Bungalows Ltd. v. Maritime Life Assurance Co.*, [1994] 2 S.C.R. 490 (Can.), available at <http://canlii.ca/t/1frs7> (last visited Sept. 24, 2012). See also *Jones v. New York Guaranty & Indemnity Co.*, 101 U.S. 622, 628 (1879) (“A court of equity abhors forfeitures, and will not lend its aid to enforce them. Nor will it give its aid in the assertion of a mere legal right contrary to the clear equity and justice of the case.”) (internal citations omitted).

cases, tweaking the remedies encourages wider use of orphan works (by reducing the user's risk), and will also narrow the scope of the problem by encouraging owners to make themselves locatable. However, in projects that involve a large number of potential claimants, such as mass digitization projects, the aggregation of many small claims through a class action may still pose a significant risk of excessive damage awards.⁶⁴ Statutory reform to eliminate, or at least reduce, the threat of a class action filed on behalf of owners of orphan works might seem desirable as well.

It is possible, however, that accepting the rationale that justifies remedy tweaking may also reduce the risk of class action suits without any need for statutory reform: recognizing the relevance of the owner's locatability may prevent a court from certifying a class that indiscriminately includes owners of orphan works with owners of non-orphan works. Moreover, if the behavior of the owner, and not only that of the user, becomes part of the calculus, then it might not even be possible to certify two subclasses (non-orphan works and orphan works) because each of such subclasses may not exhibit a sufficient degree of commonality. It might be possible, in theory, to overcome this hurdle by defining a large number of sub-classes, but at some point, the case may simply not be suitable for a class action. And practically, even before reaching this point, such a case might be much less attractive for class action lawyers and plaintiffs, because the larger the number of subclasses, the higher the cost of litigating such a case, and the lower the expected payoff to the lawyer and the representative plaintiffs of each subclass.

VI. JUSTIFYING REMEDY TWEAKING

Part V, *supra*, explained how remedy tweaking could provide a common-law solution to the orphan works problem by reducing the risk resulting from using such works, and encouraging owners who wish to be able to entitle to the full range of remedies, to take steps to prevent their works from becoming orphaned. This Part explains why this solution is not only useful but also justified. It begins by probing deeper into the orphan works problem, explains why remedy tweaking addresses the root causes of the problem, and proceeds to show how this solution is not a radical one, but in fact a common solution to similar and well known problems in various areas of the law, such as contract law, tort, and property.

64. See Samuelson & Wheatland, *supra* note 62, at 490; see also Bracha, *supra* note 12, at 1828 (illustrating how large aggregate damage awards can get when each violation entails an award between \$750 and \$30,000).

A. WHY DO WORKS BECOME ORPHANS?

Understanding the root causes of the orphan works problem is a useful step towards finding appropriate solutions. The problem tends to be described as a demand side problem: a problem of users who refrain from using works because they cannot locate the owners to seek permission and are deterred by the risk of liability. But focusing on the demand side and ignoring the supply side neglects that the orphan works problem is one of copyright owners' own making. This neglect limits our understanding of the problem and, therefore, tends to lead to proposed solutions that are equally focused on the demand side and ignore solutions that take into account the owners' part in creating the problem.⁶⁵

Shifting attention to the supply side raises an immediate puzzle. If there is demand for reusing orphan works, then one could expect that rational owners would have an incentive to make themselves easily locatable. After all, when there is demand, the forgone uses are lost opportunities for copyright owners, not only for users and the public.⁶⁶ This could suggest that the orphan works problem is not a serious one, or that the problem is not one that we—as society—should be concerned about. Arguably, if works are not used because it is too costly to locate the owner, using them is actually inefficient because the cost outweighs the benefits. The relevant cost is not only the marginal cost of reproduction (which the user is willing to incur) but also the transaction costs that a society that has chosen to have a copyright system must be willing to accept.

Furthermore, it can be expected that digitization and the Internet will probably continue to create new ways to exploit works, while also reducing some of the costs of finding owners and communicating with them. The combination of higher demand and lower transaction cost will inevitably cause the orphan work problem to shrink. In other words, it is possible that the brunt of the orphan works problem is felt only because the speed with which digitization and the Internet have created new demand has not *yet* been met with an equally speedy supply, but it will. The problem, accordingly, might be seen as a temporary one, and the policy implication is that we only need some patience, not reform.

While there may be some merit to this point of view, it is only partly correct because the orphan work problem is not a temporary one—it is endemic to copyright. The orphan works problem runs deeper because it reflects discrepancies between the social benefits of using works and the

65. *See infra* Section VII.A.

66. *Id.*

private benefits accruing to the owner, and between the social costs and private costs of forgone uses.⁶⁷ This mismatch between public and private costs and benefits arises because a copyright owner has to incur some expenses in order to maintain herself locatable. Therefore, the owner will rationally incur those costs only if the present value of the benefits that she expects to gain is larger than the costs that she needs to incur.⁶⁸

The owner, however, cannot expect to capture the full social benefit from the use of her work, only a portion of it.⁶⁹ In addition, whatever license fees she might expect to earn in the future is discounted because the future is remote and unknown.⁷⁰ Yet, the decision of how much to invest in maintaining herself locatable depends only on the ratio between the owner's private benefits and costs. Although the public benefits from future use, the public costs of forgone uses are not internalized by the owner. This mismatch between the expected social benefit and the expected private benefit to the owner leads owners to make socially sub-optimal investments in maintaining themselves locatable. In other words, owners do not internalize the full cost of their works becoming orphan, and the current system does not provide them enough incentive to invest in making themselves locatable.⁷¹

A similar problem exists on the user side. In many cases the social benefit of using the work exceeds the benefit to the user. This limits the amount of resources that the user is willing to spend in order to locate the owner and pay license fees to legally use the work. Under a strict permission-first rule,

67. See Varian, *supra* note 56, at 970–71 (recognizing the mismatch between the private costs and benefits for owners and users and the social optimum).

68. STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 87 (2004) (noting that bargaining may not occur when mutually beneficial agreements exist when the costs of bargaining—including the costs of coming together and the time and effort devoted to the bargaining process itself—outweigh the expected benefits).

69. Cf. Bracha, *supra* note 12, at 1838 (emphasizing that users cannot appropriate the full social value of their uses and therefore will abandon projects when the costs of getting permission are too high); see also Brett M. Frischmann & Mark A. Lemley, *Spillovers*, 107 COLUM. L. REV. 257, 259–61 (2007) (defining the concept of “spillovers,” or positive externalities, and explaining why owners of intellectual property never fully capture the full value of their creations).

70. The magnitude of this discount may vary among different owners. As Randy Picker notes, some owners might actually be overly optimistic about the prospects and value of future uses. Randal Picker, Panel Discussion, *Private Digital Libraries and Orphan Works* at the Berkeley Symposium: “Orphan Works and Mass Digitalization: Obstacles and Opportunities” (2012), available at <http://www.law.berkeley.edu/11731.htm> This phenomenon is also consistent with Sprigman and Buccafosco's “Creativity Effect.” Christopher Sprigman & Christopher Buccafosco, *The Creativity Effect*, 78 U. CHI. L. REV. 31 (2011).

71. Varian, *supra* note 56, at 971.

using an orphan work becomes too risky and therefore too costly from the user's perspective, even when it is socially desirable to use it. Consequently, users may make socially sub-optimal investment in locating owners.⁷²

B. LEAST-COST AVOIDERS AND LEAST-COST ENABLERS

Looking at the orphan works problem from both the demand and supply sides and restating it as a problem of costs (and benefits) that are not fully internalized by the relevant parties reveals that the orphan works problem is quite familiar. The problem ceases to be seen as an idiosyncratic copyright anomaly and instead appears to reflect a common issue that exists throughout the law; and one that has familiar solutions.⁷³

Restated as a problem of costs that are not fully internalized by the copyright owner or the user, we can approach a solution by invoking the concept of the *least-cost avoider* from the law and economics literature on tort law. Beginning with Coase's recognition that harm is often reciprocal in nature⁷⁴—caused by injurers as well as victims—the notion of the least-cost avoider supports the claim that the party who can avoid the harm at the lower cost should have a duty to take measures to avoid such harm.⁷⁵ If the same logic is applied to the orphan works problem, the question becomes which of the parties is capable of avoiding the orphan works problem and the resulting forgone socially valuable uses at a lower cost.⁷⁶ In the case of orphan works, the copyright owner is typically the least-cost avoider because she is the one who always knows who she is and where she can be found.⁷⁷ Typically, no one else knows that better than her.⁷⁸

The analogy to accidents and to the notion of the least-cost avoider is useful but incomplete. In the context of accidents, the concern is with avoiding harm that occurs in the course of involuntary relationships between injurers and victims who cannot reach a mutually beneficial agreement about

72. *Id.*

73. *Id.* at 971–72.

74. See R. H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 2 (1960).

75. See GUIDO CALABRESI, *THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS* 135–40 (1970).

76. See Dorfman & Jacob, *supra* note 15, at 94 (“[T]he breach of copyright should be assessed by reference to a more relaxed set of considerations familiar in the law of torts: who the cheapest cost avoider is; who has better information and better ability to prevent the harm; who is better able to carry the final burden; how to balance between the competing claims of the parties (to freedom, security, and so on).”).

77. Bracha, *supra* note 12, at 1838–39.

78. *Id.*

their respective activities.⁷⁹ In contrast, in the copyright context we are concerned with designing legal rules that will enable voluntary transactions and their associated benefits whenever those transactions can maximize social utility. The problem is not how to deter harmful activities but how to encourage beneficial ones. Thus, the nature of the problem is the mirror image of the accident problem, and, therefore, instead of talking about *least-cost avoiders*, it may be more accurate to talk about *least-cost enablers*. This requires us to unpack the transaction and identify, with respect to each of its elements, the party that can enable it at a lower cost. It may take one party to avoid an accident,⁸⁰ but for a voluntary licensing transaction to take place both a user and a copyright owner are necessary. Therefore, each party can take measures to facilitate the transaction, and optimal legal design would encourage each party to take those measures that she is better positioned to take and act where she is the least-cost enabler. For a transaction to take place, the user must communicate his wishes to use the work—the user is the least-cost enabler of this element of the transaction. Copyright law recognizes that truth by requiring users to seek permission. But the user cannot communicate his wishes to an unlocatable owner, even though, as noted above, the owner is typically the one who always knows who she is and where she can be found. The owner is the least-cost enabler of this element of the transaction. Unfortunately, the law is not always as cognizant about this point.

However, if the copyright owner is the least-cost enabler of her part of the transaction (or the least-cost avoider of the orphan works problem) but does not have good incentives to maintain ownership information available, the law can create these incentives with sticks, carrots, or both.

One type of carrot could be to increase the owner's expected benefits by increasing the term of copyright protection, but this solution is unlikely to be effective because copyright term is already so long; the present value of any additional profit resulting from longer terms is virtually zero.⁸¹ Moreover,

79. If there were no transaction costs, injurers and victims could reach agreements maximizing the joint value of their activities without legal intervention. See Coase, *supra* note 74.

80. See SHAVELL, *supra* note 68, at 189 (“The notion of the *least-cost avoider* applies in situations in which the risk of accidents will be eliminated if *either* injurers or victims take care. In such situations it is clearly wasteful for *both* injurers and victims to take care; rather, it is optimal for the type of parties who can prevent accidents at least cost—the least-cost avoiders—alone to take care.”) (emphasis in original).

81. See Varian, *supra* note 54, at 968. See also Brief for George A. Akerlof, et al. as Amicus Curiae Supporting Petitioners at 8, *Eldred v Ashcroft*, 537 U.S. 186 (No. 01-618), available at <http://eol.law.harvard.edu/openlaw/eldredvashcroft/supct/amici/economists.pdf>.

since copyright entails social costs, this solution will likely be counterproductive. Indeed, the orphan works problem itself is partly a result of exceedingly long copyright terms and automatic protection.⁸² Another type of carrot could be to subsidize the cost of maintaining ownership information by creating publicly funded registries. The utility of this solution is limited because even when registries are available, in order to be effective, owners still need to have their works registered and later record any changes in ownership or in their contact details. Therefore, to be effective, the carrot of a publicly funded registry may need to be accompanied with some sort of stick in the form of a negative consequence for failure to register.⁸³ The fact that the orphan works problem persists in countries that have federally funded copyright registries, such as the United States or Canada, illustrates this point.⁸⁴

Given the limitations on using plain carrots as incentives to maintain ownership information, the incentive must come in the form of a stick. Remedy tweaking provides this kind of stick. This solution recognizes that the social cost of orphan works may be avoided, and the least-cost avoider is generally the copyright owner. Therefore, a copyright system that seeks to maximize the voluntary use of works would seek to make sure that owners have proper incentives to keep themselves locatable. Remedy tweaking recognizes that while the copyright owner's exclusive right entails a corresponding duty on non-owners to refrain from copying the work without permission, the copyright owner has a duty to make herself locatable.

C. STICK OR CARROT?

Even though remedy tweaking can be formulated as a stick (“if you don’t make yourself easily locatable *ex ante* you will be entitled to less *ex post*”), it can also be formulated as a carrot (“you, the owner, can choose between investing today in order to keep yourself locatable in the future, in which case you will be entitled to all the remedies, OR you can choose to invest less today, but might earn less in the future”). Formulating remedy tweaking as a carrot is not merely a semantic point or a rhetorical device. In fact, relaxing the permission-first rule by recognizing a duty to be locatable does not necessarily harm copyright owners, and in fact may be beneficial to them. To see why, recall that the decision to orphan the work is based on the copyright

82. Varian, *supra* note 54, at 968–69; LAWRENCE LESSIG, FREE CULTURE: HOW BIG MEDIA USES TECHNOLOGY AND THE LAW TO LOCK DOWN CULTURE AND CONTROL CREATIVITY 135 (2004).

83. *See* Van Houweling, *supra* note 52.

84. *See* Bracha, *supra* note 11, at 1829 (noting the lack of sufficient incentive to register copyright).

owner's calculation of the present value of future uses, which in turn, depends on the copyright owner's prediction of what the future might entail.⁸⁵ But the owner's ability to predict who might wish to exploit her work in the future and for what purpose is inherently limited. In contrast, the future user who wants to use the work knows who she is, and may also have better information than the owner about the expected value of the use.⁸⁶ Unfortunately, under a strict permission-first rule the owner's decision not to invest in keeping herself locatable may be rational, but is also partly irreversible, because strict liability might deter users from using it and thus sending the owner a signal that there might be renewed interest in the work. To prevent this from happening, some owners may actually over-invest in keeping themselves locatable. Remedy tweaking might therefore benefit owners because it encourages users to reuse the work, and thereby signal to owners that there is renewed interest in their works even if it has become orphaned. Rather than incurring the full cost of maintaining themselves locatable, a relaxed rule allows the owner to economize on the information that users provide by using.

While remedy tweaking may reduce the owner's bargaining power vis-à-vis the user, this may be a small price to pay because it reduces over deterrence and opens up new opportunities for compensable transaction. The alternative—under a strict permission-first rule—can make both the owner and the user worse off. In any event, an owner who wishes to retain a greater bargaining power can do that by maintaining herself locatable.⁸⁷ Therefore, remedy tweaking allows both owners and users to choose between different strategies with respect to managing opportunities and risks.

D. "SIT BACK, DO NOTHING" IS RARELY THE LAW

As discussed in Part III, *supra*, Judge Chin's statement that copyright law allows owners to sit back and do nothing does not fully reflect current copyright law. Moreover, if understood to be more than merely a description of an entitlement protected by a property rule, rather than a starting point for the study of the instances in which the law may deviate from it, the statement is inconsistent with the law more broadly. If the statement stands for the proposition that a copyright owner may sit back, do nothing, regardless of the circumstances and regardless of any consequences, this would be a rare occurrence within our legal system. As the following examples illustrate, the law rarely recognizes such an unfettered right to sit back and do nothing, and

85. See *supra* text accompanying note 70.

86. See *supra* text accompanying note 12.

87. See *supra* Part I.

frequently demands or encourages right holders to take action to minimize negative consequences arising from their inaction.

This Article has already alluded to tort law and the concept of the least-cost avoider,⁸⁸ but tort law can teach more than that. While tort law may protect victims' rights over their bodies or property, it protects those entitlements under two distinct regimes, which can be roughly described as a non-accident regime and an accident regime, the differences between which are relevant to the current discussion. As Dorfman and Jacob note:

The “non-accident” regime, as generally manifested in the intentional torts and their strict liability approach, carries with it a “package” ensuring that the wrongdoer will seek the consent of the right-holder and, thus, engage the latter in a market interaction. If one does not seek permission prior to entrance—she will not only have to pay for the actual damage, but also be exposed to criminal liability, punitive damages, disgorgement of profits, etc. The accident regime, on the other hand, assumes that market transaction is not a viable option (whatever the reason may be). At the same time, it aims to promote productive activities, at least to a certain extent. Therefore, the accident regime will allow actors to inflict damage on others: Negligence law would only require them to act in a reasonable manner, while absolute liability would impose on them the entire risk of the materialized accident, regardless of reasonable care being taken.⁸⁹

When tort law protects the owner of a right under an accident regime, rarely will it permit the owner to sit back and do nothing if it would be reasonable to take action to minimize risk and maximize social utility. That is, strict liability rules, completely disregarding the victim's behavior, are a rarity. As Shavell notes, “strict liability does not lead to the socially optimal outcome for the obvious reason that it fails to furnish victims a motive to take care.”⁹⁰ Even when the law adopts a standard of strict liability, the victim's behavior may be relevant, either by recognizing a defense of contributory negligence or by looking at the parties' comparative negligence.⁹¹ Moreover, even under the non-accident regime—as in the case of intentional torts, where invasions to one's tangible property are actionable as trespass and impose strict liability on the tortfeasor—the choice of this particular form of liability is “not a logical feature of ownership or private property . . . [but] to an important extent a feature of certain facts about the

88. *See supra* Section VI.B.

89. Dorfman & Jacob, *supra* note 15, at 83.

90. SHAVELL, *supra* note 68, at 188.

91. *Id.*

world occupied by owners and non-owners of material objects.”⁹² Strict liability for intentional torts is justified because those torts “involve not a conflict between legitimate (productive) activities but a coerced transfer of wealth to the defendant in a setting of low transaction cost.”⁹³ Therefore, “[s]o long as there is a market, the law will null all incentives to bypass it.”⁹⁴ Yet the very considerations that justify strict liability when market transactions are available warrant a switch to an accident-like regime where the injurer’s activity, while socially productive, may inflict harm on another, yet the costs of transacting are prohibitively high.⁹⁵ In such cases, “tort law will provide . . . ‘guidelines’ about how to act with due care. In doing so, courts and legislators take into account not only the interest of the injurer and the injured party . . . but also society’s, attempting to maximize the overall welfare.”⁹⁶

Furthermore, tort law encourages action not only in order to minimize the occurrence of accidents, but also to minimize harm arising from accidents that do occur. Thus, injured victims who fail to take reasonable steps to minimize their harm after accidents occur, for example by seeking post-injury medical treatment, may not be entitled to compensation for their full damages. In reaching this conclusion, the Supreme Court of Canada reasoned that this “‘duty to mitigate’ derives from the general proposition that a plaintiff cannot recover from the defendant damages which he himself could have avoided by the taking of reasonable steps.”⁹⁷

92. See Dorfman & Jacob, *supra* note 15, at 79.

93. *Id.* at 85 (quoting RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 205 (1999)).

94. *Id.*

95. See *id.*

96. *Id.* at 86. Dorfman and Jacob describe several features of copyright, which, according to them, justify greater reliance of accident-like rules of liability. They claim that knowing the identity of the owner of tangible property is typically easier than ascertaining the owner of a copyright, and they use orphan works as an example: “We all know the term ‘orphan work,’ but in tangible property this term is not yet to be coined. The term ‘orphan house’ has an entirely different meaning.” *Id.* at 92.

97. *Janiak v. Ippolito*, [1985] 1 S.C.R. 146, 36 (Can.); see also *RESTATEMENT (SECOND) OF TORTS* § 918 (1965). Section 918 of the Restatement (Second) of Torts provides for the following:

(1) Except as stated in Subsection (2), one injured by the tort of another is not entitled to recover damages for any harm that he could have avoided by the use of reasonable effort or expenditure after the commission of the tort. (2) One is not prevented from recovering damages for a particular harm resulting from a tort if the tortfeasor intended the harm or was aware of it and was recklessly disregarding of it, unless the injured person with knowledge of the danger of the harm intentionally or heedlessly failed to protect his own interests.

A duty to mitigate is also a common feature of contract law. The doctrine of mitigation holds that a victim of breach of contract may not be entitled to damages that he could avoid by taking reasonable steps.⁹⁸ It is “a doctrine based on fairness and common sense, which seeks to do justice between the parties in the particular circumstances of the case.”⁹⁹ Although not frequently invoked in the area of intellectual property law, some courts were willing to apply the doctrine of mitigation in the area. In a recent patent case, the court held that “it is entirely appropriate for a defendant to assert a defense of failure to mitigate damages when considering what amount of compensation is appropriate,”¹⁰⁰ and in an earlier copyright case, the court held that “[t]he mitigation doctrine of avoidable consequences is a fundamental rule of damages requiring the injured party to take advantage of reasonable opportunities to minimize his damages and avoid or prevent loss.”¹⁰¹

Examples of duties to take action—failure of which may result in losing one’s entitlement to a stronger relief—may be found not only in tort and contract law, but also in property law, and in rules governing the legal process. Adverse possession is the most obvious example in property law,¹⁰² but numerous other examples rejecting a “sit back, do nothing” approach exist, including a landowner’s potential liability for injuries inflicted on

RESTATEMENT (SECOND) OF TORTS § 918.

98. *Brit. Westinghouse Elec. & Mfg. Co. v. Underground Elec. Rys. Co. of London*, [1912] 1 A.C. 673 (H.L.) 689 (“The fundamental basis is thus compensation for pecuniary loss naturally flowing from the breach; but this first principle is qualified by a second, which imposes on a plaintiff the duty of taking all reasonable steps to mitigate the loss consequent on the breach, and debars him from claiming any part of the damage which is due to his neglect to take such steps.”).

99. *See Southcott Estates Inc. v. Toronto Catholic Dist. Sch. Bd.*, [2012] SCC 51, ¶ 25 (Can.).

100. *IMX, Inc. v. E-Loan, Inc.*, 748 F. Supp. 2d 1354, 1361 (S.D. Fla. 2010) (rejecting the plaintiff’s contention that a mitigation of damages defense does not exist in patent law and noting that although “[s]uch a defense may rarely be relevant in a patent infringement case . . . it is not inappropriate”).

101. *Gener-Villar v. Adcom Group, Inc.*, 560 F. Supp. 2d 112, 133 (D.P.R. 2008). However, the court there found that the plaintiff sufficiently mitigated his damage by objecting to the defendant’s acts. *Id.* at 134.

102. *See Lydia Loren, Abandoning the Orphans: An Open Access Approach to Hostage Works*, 27 BERKELEY TECH. L.J. 1431 (2012); *see also* Aryeh L. Pomerantz, *Obtaining Copyright Licenses by Prescriptive Easement: A Solution to the Orphan Works Problem*, 50 JURIMETRICS 195, 204–05 (2010); Matthew W. Turetzky, *Applying Copyright Abandonment in the Digital Age*, 2010 DUKE L. & TECH. REV. 19, 24 (2010).

trespassers, notice requirements,¹⁰³ and the doctrines of abandonment¹⁰⁴ and prescriptive easement.¹⁰⁵

In civil litigation, duties to take action underlie limitation periods,¹⁰⁶ or the doctrine of laches.¹⁰⁷ They encourage litigants to act sooner rather than later to vindicate their rights, and thereby reduce costs of uncertainty, opportunism, and protracted litigation.¹⁰⁸ Rules of civil procedure are also quite averse to a problem that we may describe here as “orphan defendants.” Even though normally a plaintiff needs to serve his complaint on the defendant personally,¹⁰⁹ the fact that a defendant is unlocatable does not necessarily bar the court from issuing a judgment against an unlocatable defendant where an alternative to personal service is permitted.¹¹⁰ If personal or alternative service is impractical, a court may make an order for substituted service¹¹¹ or “where necessary in the interest of justice, may dispense with service.”¹¹² If the defendant emerges she may have some limited opportunity to set aside a default judgment “on such terms as are just,”¹¹³ but otherwise sitting back and doing nothing may lead to serious consequences to the orphan defendant.

The point here is not to argue that each of these rules maps directly onto the orphan works problem and provides a ready-made doctrinal solution. Obviously many of them, developed in other contexts, may not immediately apply, or may require various adaptations. Instead, the point is to emphasize what many of these rules have in common: a recognition that the law often discourages a sit back, do nothing approach and often encourages least-costs avoiders to take action by creating consequences for inaction.

103. See Herbert J. Hovenkamp, *Notice and Patent Remedies*, 88 TEXAS L. REV. 221, 224–27 (2011), available at <http://texasrev.com/sites/default/files/seealso/vol88/pdf/88TexasLRevSeeAlso221.pdf>.

104. See Loren, *supra* note 102; Turetzky, *supra* note 102.

105. See Loren, *supra* note 102; Pomerantz, *supra* note 102, at 220–22.

106. See, e.g., Limitations Act, S.O. 2002, c. 24, s. B (Can.), available at http://www.e-laws.gov.on.ca/html/statutes/english/elaws_statutes_02l24_e.htm (last visited June 28, 2012).

107. See *M.(K.) v. M.(H.)*, [1992] 3 S.C.R. 6 (Can.).

108. *Id.*

109. See, e.g., Ontario Rules of Civil Procedure, CANLII, RRO 1990, Reg 194, s. 16.02, <http://www.canlii.org/en/on/laws/regu/rro-1990-reg-194/latest/rro-1990-reg-194.html> (last visited May 8, 2012).

110. *Id.* s. 16.03.

111. *Id.* s. 16.04.

112. *Id.*

113. *Id.* s. 19.08.

E. JUDICIAL AUTHORITY TO TWEAK REMEDIES

A possible objection to remedy tweaking is that even if remedy tweaking is desirable, courts may not have the power to tweak statutory remedies unless the statute permits them to do that. Thus, a distinction might be drawn between discretionary remedies under §§ 502 (injunctions)¹¹⁴ and 503 (“[i]mpounding and disposition of infringing articles”),¹¹⁵ which the court *may* order, and remedies under § 504 (actual damages, the infringer’s profit, and statutory damages) to which the copyright owner is statutorily “entitled.”¹¹⁶ Therefore, the question becomes whether courts can apply common law principles in determining remedies under § 504. A full exploration of this question is beyond the scope of this Article, and if the answer is no, then this Article should be read as providing reasons for statutory reform. However, the answer may be yes for at least the following reasons:

First, it is an established principle of statutory construction that “where Congress uses a common-law term in a statute, we assume the term comes with a common law meaning, absent anything pointing another way.”¹¹⁷ This would suggest that the copyright owner’s entitlement to damages and accounting for profits under § 504 ought to be subject to established principles applying to those remedies under the common law. These may include doctrines that limit those remedies, such as mitigation. As noted above, some courts have taken it for granted that the principle of mitigation applies to damages for infringement of intellectual property.¹¹⁸

Similarly, despite the choice of the word “entitled,” at least one court has held that accounting for profit “had been given in accordance with the principles governing equity jurisdiction, not to inflict punishment but to

114. 17 U.S.C. § 502(a) (2010) (“Any court having jurisdiction of a civil action arising under this title may . . . grant temporary and final injunctions on such terms as it may deem reasonable to prevent or restrain infringement of a copyright”); *see also* eBay Inc. v. MercExchange, 547 U.S. 388, 392–93 (2006) (confirming the discretionary nature of injunctions).

115. 17 U.S.C. § 503(a)(1) (“At any time while an action under this title is pending, the court may order the impounding, on such terms as it may deem reasonable”); 17 U.S.C. § 503(b) (“As part of a final judgment or decree, the court may order the destruction or other reasonable disposition”).

116. 17 U.S.C. § 504(b) (2010) (“The copyright owner is entitled to recover the actual damages suffered by him or her as a result of the infringement, and any profits of the infringer that are attributable to the infringement”).

117. *See* Microsoft Corp. v. i4i Ltd., 131 S. Ct. 2238, 2245 (2011) (internal quotations and citations omitted).

118. *See supra* notes 100, 101 and accompanying text.

prevent an unjust enrichment.”¹¹⁹ This should justify denying (or reducing) the award of profits when the rationales underlying this remedy do not exist. In *Taylor v. Meirick*,¹²⁰ Judge Posner explained that the purpose of this remedy is to discourage infringement and encourage users to seek permission and facilitate voluntary exchange.¹²¹ This rationale does not hold when it is prohibitively costly to seek permission and voluntary exchange is impossible. Another rationale of preventing unjust enrichment is “to prevent the infringer from unfairly benefiting from a wrongful act.”¹²² While a user of an orphan work may benefit from a wrongful act, it is questionable whether she benefits from it “unfairly.” When seeking permission would be unreasonable under the circumstances and the only alternative to the unauthorized use is not to use at all—an alternative that leaves the user and society worse off, and the owner no better off—it is quite difficult to characterize the act as unfair.

Second, and consistent with the first point, courts have not hesitated to apply common law principles of liability in copyright cases even when they have no statutory basis. For example, in *Sony Corp. of America v. Universal City Studios*,¹²³ the Court held that:

The absence of . . . express language in the copyright statute [that renders a person liable for the infringing acts of another] does not preclude the imposition of liability for copyright infringements on certain parties who have not themselves engaged in the infringing activity. For vicarious liability is imposed in virtually all areas of the

119. See *Sammons v. Colonial Press*, 126 F.2d 341, 346 (1st Cir. 1942) (citing *Sheldon v. Metro-Goldwyn Pictures Corp.*, 309 U.S. 390, 399 (1940)). Even though these cases were decided under the 1909 Copyright Act, there is nothing to suggest that Congress changed the nature of this remedy when it enacted the present Act in 1976. See Kenneth E. Burdon, *Accounting for Profits in a Copyright Infringement Action: A Restitutionary Perspective*, 87 B.U. L. REV. 255, 270 (2007) (explaining that the “restitutionary theory” of the 1909 Act remained unchanged after the Act’s revision in 1976).

120. *Taylor v. Meirick*, 712 F.2d 1112 (7th Cir. 1983).

121. More specifically, Judge Posner noted the following:

By preventing infringers from obtaining any net profit it makes any would-be infringer negotiate directly with the owner of a copyright that he wants to use, rather than bypass the market by stealing the copyright and forcing the owner to seek compensation from the courts for his loss. Since the infringer’s gain might exceed the owner’s loss, especially as loss is measured by a court, limiting damages to that loss would not effectively deter this kind of forced exchange.

Id. at 1120.

122. H.R. Rep. No. 94-1476, at 161 (1976), *reprinted in* 1976 U.S.C.C.A.N. 5659, 5777; *see also* Burdon, *supra* note 119, at 267–71.

123. 464 U.S. 417 (1984).

law, and the concept of contributory infringement is merely a species of the broader problem of identifying the circumstances in which it is just to hold one individual accountable for the actions of another.¹²⁴

Absent an express statutory mandate, it would be a strange outcome if general principles and concepts of liability that exist throughout the law could only be used to expand the scope of liability when justified, but not to contract it when appropriate. Such an outcome is also inconsistent with earlier and current case law. For example, in *Lawrence v. Dana*, the court noted that even when copying exceeds the limits of fair use, “cases frequently arise in which, though there is some injury, yet equity will not interpose by injunction to prevent the further use, as where the amount copied is small and of little value, if there is no proof of bad motive, or where there is well-founded doubt as to the legal title, or where there has been long acquiescence in the infringement, or culpable laches and negligence in seeking redress, especially if it appear that the delay has misled the respondent.”¹²⁵ Very recently, the Ninth Circuit barred copyright infringement claims on the basis of the equitable doctrine of laches.¹²⁶

In sum, remedy tweaking can provide a common law solution to the orphan works problem. First, remedy tweaking recognizes that the orphan works problem is a supply-side problem just as it is a demand-side problem. It is not only a problem of users unable to locate owners, but of owners’ suboptimal investment in maintaining themselves locatable. A solution for the orphan works problem that seeks to preserve copyright’s market orientation needs to take into account the incentives that users and owners face and allocate the cost of maintaining a market system between users and owners.

Second, remedy tweaking takes into account the information that each of the user and the owner possesses. Remedy tweaking assumes that the user has the best information about his contemplated use and therefore should take reasonable steps to communicate his plans to the owner by seeking permission. The owner, on the other hand, has the best information about her identity and how she can be reached, and therefore should take reasonable steps to make this information available to potential users. A strict permission-first rule, which ignores the owner’s ability to take steps to maintain themselves locatable, puts the entire risk on users and leads to

124. *Id.* at 435 (internal citations omitted).

125. 4 Cliff. 1, 61 (C.C. Mass. 1869).

126. *See* *Evergreen Safety Council v. RSA Network Inc.*, 697 F.3d 1221 (9th Cir. 2012); *Petrella v. Metro-Goldwyn-Mayer, Inc.*, 695 F.3d 946 (9th Cir. 2012).

suboptimal utilization of works. By recognizing that copyright owners should have a duty to maintain themselves locatable and by tweaking the remedies available to them if they fail to fulfill this duty, courts can adjust the user's duty to seek ownership information and permission with a corresponding duty of the owner to provide such information. Doing that will shift some of the risk of using orphan works from the user to the owner, encourage owners to take measures to reduce the extent of the problem, and encourage greater utilization of works.

Finally, remedy tweaking leads to an important conclusion: while it may be convenient to talk about an "orphan works problem," it may be more accurate to recognize that we are indeed facing a more general problem of how to calibrate owners' and users' correlative rights and duties over a continuum and in a wide range of circumstances. In many cases "permission first" is a sensible rule, and when it is sensible, copyright owner can sit back and do nothing. Yet in other cases, a copyright system, like any other system of private ordering, cannot function effectively without imposing some duties on owners.¹²⁷ It would not be an anomaly, where appropriate, to require copyright owners to maintain information about how they can be reached, to send take-down notices when they do not want their works to be used in a particular way, or to adopt simple and effective measures to communicate their wishes.

With these observations and conclusions, this Article now examines and critiques other solutions to the orphan works problem.

VII. QUASI-DOGMATIC AND DOGMATIC SOLUTIONS

This Part canvasses and critiques current proposed solutions to the orphan works problem, which can be viewed as "dogmatic." To some degree or another, all of them ignore the fact that the orphan works problem is of the copyright owners' own making, and instead focus exclusively on the users. This Part begins with a discussion of the U.S. Copyright Office proposal from 2006, which despite being similar to this Article's proposed remedy-tweaking solution, contains some elements that place it on the dogmatic side. Next this Part discusses Canada's statutory orphan works regime, as well as collective licensing and extended collective licensing.

127. See generally Lisa M. Austin, *Possession and the Distractions of Philosophy*, in PHILOSOPHICAL FOUNDATIONS OF PROPERTY LAW, (J.E. Penner & Henry Smith, eds., forthcoming, 2013), available at <http://ssrn.com/abstract=2062289> (last visited Nov. 4, 2012) (arguing that every system of private ownership requires a method of allocating title and that this must conform to the rule of law demand that it be sufficiently public).

A. THE U.S. COPYRIGHT OFFICE REMEDY TWEAKING PROPOSAL

Following extensive examination,¹²⁸ in 2006 the Copyright Office issued its Orphan Works Report (“Report”).¹²⁹ The Report recommended statutory reform to limit the scope of remedies available against defendants who were unable to locate the copyright owner of a work after a “good faith, reasonably diligent search.”¹³⁰ The proposed statutory language is highly similar to this Article’s proposal, and, if implemented, could be used as a model for this Article’s proposal as well; there are many uses where there would be hardly any difference between the Copyright Office’s proposal and the proposal in this Article. However, despite the similarities, this Section identifies some differences between the Report and this Article’s proposal that reflect the copyright dogma’s influence on the Report.

The first major difference is that the Report does not recognize copyright owners’ contribution to the problem and their ability to minimize it. While it explains in detail why works become orphans and the problems of identifying and locating copyright owners,¹³¹ the Report fails to acknowledge the fact that the orphan works problem is of copyright owners’ own making, and that generally, copyright owners are the problem’s least-cost avoiders. The only instance where the Report seems to acknowledge the owner’s ability to avoid the problem is when it discussed whether to adopt a formal approach or a case-by-case approach.¹³² It mentioned an argument made by Save the Music & Creative Commons that a formal approach is justified because “it places the burden of keeping contact information current on the party best able to bear it: the owner of the copyright.”¹³³

Eventually, the Report adopted a case-by-case approach. This approach attracted support for two different reasons. One was that “the definition of a reasonable search will vary greatly depending on the category of the work, and it will be impossible to prescribe *ex ante* the searches that will be

128. See David R. Hansen, *Orphan Works: Mapping the Possible Solution Spaces 2* (Berkeley Digital Library Copyright Project, White Paper No. 2, 2012), <http://ssrn.com/paper=2019121> (last visited June 23, 2012).

129. See *supra* note 49.

130. *Id.* at 127.

131. *Id.* at 21–29.

132. *Id.* at 72 n.217 (“Under a ‘case-by-case’ standard, the user’s search is judged on a case-by-case basis, *i.e.* the adequacy of a search for an owner is considered *ad hoc*, with reference to the circumstances prevailing at the time of the search. Under the ‘formal’ approach, there is a pre-set list of required searches, and once the user performs those searches without success, the work is deemed orphaned.”).

133. *Id.* at 72.

adequate for each category.¹³⁴ To the extent that the Report adopted the case-by-case approach for its flexibility and adaptability, the choice is consistent with my proposal,¹³⁵ and it is possible that despite the lack of explicit recognition of copyright owners' contribution to the problem, courts would take that into account when deciding whether the user's search was reasonable. Indeed, as Professor Lessig suggested, courts could interpret the limitation on remedies as an "implicit recognition that copyright owners have a responsibility to help make the copyright system function more efficiently . . . [and] that the copyright owner bears some burden (the burden of maintaining accessibility) as a condition of getting the full benefits of copyright law's protection."¹³⁶

However, the Report also objected to the formal approach on the grounds that it would constitute "a dramatic reversal of the current copyright regime, in which an owner need not perform any positive act to preserve his rights."¹³⁷ This statement, of course, is a clear articulation of the dogma. Therefore, since neither the Report nor the proposed statutory language explicitly recognize that copyright owners should bear some of the burden of making the system work more efficiently, there is no guarantee that judges would follow Professor Lessig¹³⁸ and recognize this implicit duty. Instead, like Judge Chin, they may resort to a dogmatic impulse, refuse to recognize such a duty without clear mandate from Congress, and decline to interpret the proposal as providing them such mandate. In fact, they might even conclude that the proposal consciously declined to recognize such duty.

Consider the following example: as noted in the Report, "if a well-known and comprehensive voluntary registry of owners existed in a particular industry or for a particular type of work, presumably any 'reasonable search' for the owner would include a search of that registry."¹³⁹ It is not clear, however, what the implications of this statement are. The statement may suggest that consulting such registry is a necessary condition in the sense that if such a registry exists and consulting it would allow the user to locate the

134. *Id.*

135. *Id.* at 113 ("We believe the truly 'ad hoc' system—where users simply conduct a reasonable search and then commence use, without formality—is the most efficient way to proceed.")

136. See LAWRENCE LESSIG, LETTER FROM STANFORD LAW PROFESSOR LAWRENCE LESSIG TO THE HON. ZOE LOFGREN (2006), available at <http://www.lessig.org/blog/archives/20060306-lofgren.pdf>.

137. REGISTER OF COPYRIGHTS, *supra* note 49, at 72 (citing the objection of the National Music Publishers' Association, Inc. & The Harry Fox Agency, Inc.).

138. See LESSIG, *supra* note 136.

139. *Id.* at 72 n.217.

owner, a user who fails to search the registry would not be eligible for any limitation on the remedy. But is consulting the registry a sufficient condition, in the sense that a user who searched the registry in vain would be entitled, as a matter of law, to a limitation on the remedy? Holding that consulting the registry would be a sufficient condition implicitly recognizes that an owner has a duty to register his works in that registry, and that failing to comply with this duty would limit the remedies available to him. Since the Report discusses only what is reasonable for users to do and falls short of explicitly mentioning any corresponding duties on owners, it may be perfectly open to a court to conclude that consulting the registry may be a necessary condition but not a sufficient one, and that since the owner has no duty to register, a diligent user should have taken additional measures to search the registry.

More generally, because the Copyright Office proposal overlooks the owners' side, it provides an ad hoc solution to a specific problem without explaining how it fits within the general juridical order of which copyright law is a part. As a result, judges may not feel confident to apply the proposal liberally and instead may treat it as an exception that should be interpreted narrowly.

Another difference between the Report and this Article's proposal, which also stems from ignoring the owners' contribution to the problem, is that the Copyright Office proposal encourages a tendency to treat orphan works as a question of status rather than a problem that exists on a continuum. The proposal explicitly refers to the "designation of a work as orphaned,"¹⁴⁰ and then turns to identify the criteria for such designation.¹⁴¹ The result is that a work either qualifies as orphaned—triggering the limitation on remedies—or does not qualify, in which case no limitation whatsoever is permitted. The result is a crude dichotomy that can produce dramatically different outcomes in similar cases.¹⁴² For example, suppose that two users contemplate similar projects. Both of them make good faith, albeit unsuccessful, efforts to locate the owner, but User A pursued one additional search, while User B did not. Now, suppose that the owner emerges and sues both of them; User A, who pursued that additional search, is able to convince the court that he was reasonably diligent, whereas User B is found to have been not diligent enough. The result is that User A may be protected from harsh remedies, while User B may be just as liable to the full panoply of remedies as if he were the worst offender who conducted no search at all. The possibility that

140. *Id.* at 71.

141. *Id.* at 71–83.

142. Admittedly, a similar problem exists under the fair-use doctrine. *See* Kozinski & Newman, *supra* note 33.

minor non-consequential differences in the facts of similar cases will result in dramatic differences in the outcome runs against deep notions of the rule of law.¹⁴³ Moreover, such outcomes are even more difficult to justify if the owner could have easily made it possible for both users to locate her but failed to do so.

Lastly, the Copyright Office's proposal contains some prerequisites that may introduce rigidity in a place where flexibility is required. For example, the limitation on the remedy may only be available when the user performed a good faith and reasonably diligent search to locate the owner prior to copying the work. While normally seeking the owner prior to using the work is a justified requirement, there are situations in which using the work while continuing to try and locate the owner does not seem objectionable.¹⁴⁴ It is easier to reach this conclusion when acknowledging that owners are at least partly responsible for the difficulty of locating them. This aspect of the Copyright Office's proposal is another reflection of its exclusive focus on the behavior of users, and the underlying influence of the dogma.

This Article's proposal avoids these problems because it does not seek to designate a status of an orphaned work. Instead, it recognizes that the locatability of the owner is a question of degree and that the concept of an orphan work is a relative one. Locating the owner of almost any work requires some effort, and very few works, if any, are truly orphaned in the sense that even with unlimited resources it would be impossible to track their owners. Therefore, within this Article's proposed framework, saying that a work is orphaned simply means that under the circumstances of a particular case, the user should not have exerted additional effort to locate the owner and that the owner's inaction justifies a limitation on the remedies. Thus, the term "orphan" is simply the conclusion of the analysis rather than a prerequisite designation.

Despite its shortcomings, the Copyright Office's proposal, had it been adopted, would have in practice allowed judges considerable latitude to implement it in a way that is consistent with my approach. While its exclusive focus on users and lack of attention to owners' contribution to the orphan work problem reflect the dogma, it is significantly better than other proposed solutions that are truly captive of the dogma, such as those discussed in the next Section.

143. Cf. David A. Strauss, *Must Like Cases Be Treated Alike?* (U. Chi. Pub. Law & Legal Theory, Working Paper No. 24, 2002), available at <http://ssrn.com/paper=312180> (last visited June 25, 2012).

144. Consider the example of the Dutch broadcasters. See Van Gompel, *supra* note 52.

B. CANADA'S ORPHAN WORKS REGIME

In 1988, as part of a major reform to its Copyright Act, Canada introduced a novel mechanism for dealing with the orphan works problem.¹⁴⁵ Section 77 established a mechanism allowing prospective users of a work to obtain a license to use it from the Copyright Board when they cannot locate the owner.¹⁴⁶ The section provides that “[w]here, on application to the Board by a person who wishes to obtain a licence to use [a published work] . . . in which copyright subsists, the Board is satisfied that the applicant has made reasonable efforts to locate the owner of the copyright and that the owner cannot be located, the Board may issue to the applicant a licence to [use the work].”¹⁴⁷ Such license is non-exclusive and is subject to such terms and conditions as the Board may establish.¹⁴⁸ If the Board issues such a license, the copyright owner may, no later than five years after the expiration of the license, collect the royalties fixed in the license or, in default of their payment, commence an action to recover them in a court of competent jurisdiction.¹⁴⁹

145. See Jeremy De Beer & Mario Bouchard, *Canada's 'Orphan Works' Regime: Unlocatable Copyright Owners and the Copyright Board*, 10 OXFORD U. COMMONWEALTH L.J. 215, 220 (2010).

146. *Copyright Act*, R.S.C. 1985, c. C-42, s. 77.

147. *Id.* s. 77(1).

148. *Id.* s. 77(2).

149. *Id.* s. 77(3). The entire section reads as follows:

77. (1) Where, on application to the Board by a person who wishes to obtain a licence to use

- (a) a published work,
- (b) a fixation of a performer's performance,
- (c) a published sound recording, or
- (d) a fixation of a communication signal

in which copyright subsists, the Board is satisfied that the applicant has made reasonable efforts to locate the owner of the copyright and that the owner cannot be located, the Board may issue to the applicant a licence to do an act mentioned in section 3, 15, 18 or 21, as the case may be.

(2) A licence issued under subsection (1) is non-exclusive and is subject to such terms and conditions as the Board may establish.

(3) The owner of a copyright may, not later than five years after the expiration of a licence issued pursuant to subsection (1) in respect of the copyright, collect the royalties fixed in the licence or, in default of their payment, commence an action to recover them in a court of competent jurisdiction.

(4) The Copyright Board may make regulations governing the issuance of licences under subsection (1).

Id. s. 77.

At first glance, this solution looks attractive, and appears to provide a pragmatic and balanced solution. This statutory mechanism recognizes that when the owner cannot be located, adhering to the permission-first rule would prevent works from being used for socially beneficial purposes, and that such an outcome is inconsistent with the purpose of the Copyright Act. Accordingly, it provides a solution by allowing the Copyright Board to issue a compulsory license that authorizes the use of a work when its owner cannot be located. However, a license will not be available unless the user made reasonable efforts to locate the owner. If the owner is locatable but refuses to negotiate or grant a license, no license will issue.¹⁵⁰ Even if the owner does not respond to any request by the user, no license will issue, as long as the owner is locatable.¹⁵¹

While the mechanism rewards the diligent user by granting a license, thereby immunizing him from any future claim by the owner should she emerge, it also considers the interests of the owner. The Board does not simply immunize the user from future liability. Rather, it fixes royalties that the users should pay, and may determine other terms and conditions. Although the owner cannot get the money as long as he is unlocatable, the mechanism provides that the owner, should he appear, and within five years after the expiry of the license, may collect the royalties from the user, and in the case of default of their payment, commence an action to recover them. Thus, rather than suing for copyright infringement (and prevailing in such a suit) and being subject to the standard three years limitation period beginning at the time of the reproduction, the owner only needs—if the user refuses to pay—to commence an action for a recovery of a debt, and can do that within a longer period, five years after the expiry of the license. The user, however, does not have to pay unless and until the owner demands payment, and therefore, in the not-unlikely event that the owner never shows up, the user does not have to pay anything while minimizing the risk of liability to zero. All in all, this mechanism seems to provide not only a pragmatic solution to the orphan works problem, but a solution that deviates as little as possible from the market mechanism, and from the bilateral relations that lie at the heart of copyright.¹⁵²

However, this solution only seems pragmatic until one considers it more carefully, and further, until one pays close attention to how the Board implements it. First, even though the statute may not strictly mandate this, the Board created a quasi-judicial procedure: not only must the user apply for

150. De Beer & Bouchard, *supra* note 145, at 226.

151. *Id.*

152. *See supra* Part III.

a license, she must also establish that the statutory requirements exist.¹⁵³ This creates an apparent anomaly: whereas scores of commercially available and highly valuable works are licensed every day through streamlined simple procedures without ever involving anything that resembles a judicial procedure (consider iTunes), and whereas only a small number of them might ever be litigated, the use of orphan works, which are almost by definition works with very little commercial value, requires a much more costly quasi-judicial procedure. This seems to be a totally inefficient use of users' private resources and of the Board's public resources.

In addition to this fundamental anomaly, the Board's implementation adds additional layers of complexity, which heighten the burden on the user. Since section 77(1) refers to a "published work,"¹⁵⁴ the Board takes the view that it could not issue a license unless the user demonstrates that the work had indeed been published. On this basis, the Board refuses to issue licenses when applicants could not provide evidence that the work had been published. In one case the Canadian Centre for Architecture in 2004 wished to reproduce and display three photographs taken in 1955, 1957, and 1967 and found at the Public Archives of Canada, for the purpose of an exhibition entitled *Les années 60: Montréal voit grand* ("The 60's: Montreal Thinks Big").¹⁵⁵ In another case the same year, the Office of the Lieutenant Governor of Québec wished to reproduce a photograph taken on the opening day of the November 1959 legislative session in a book on the history of lieutenant governors of Québec.¹⁵⁶

Second, while the Act provides that the Board must be satisfied that the applicant has made reasonable efforts to locate the owner of the copyright and that the owner cannot be located, the Act does not specify what ought to satisfy the Board.¹⁵⁷ Arguably, it would be open to the Board to adopt a very minimal threshold. For example, it could be enough for an applicant to fill an online form and click "I Agree" below a declaration that she made reasonable efforts to locate the owner and could not locate him, and if that seems too easy, a sworn declaration of the same could suffice to satisfy the Board. The Board, however, chose to require much more than that. In the

153. De Beer & Bouchard, *supra* note 145, at 226.

154. Copyright Act, R.S.C. 1985, c. C-42, s. 77(1) (Can.).

155. COPYRIGHT BOARD CANADA, 'RE CANADIAN CENTRE FOR ARCHITECTURE' 2004-UO/TI-32 (Jan. 17, 2005), <http://www.cb-cda.gc.ca/unlocatable-introuvables/other-autre/3-b.pdf>.

156. COPYRIGHT BOARD CANADA, 'RE OFFICE OF THE LIEUTENANT GOVERNOR OF QUÉBEC' 2004-UO/TI-37 (Mar. 3, 2005), <http://www.cb-cda.gc.ca/unlocatable-introuvables/other-autre/4-b.pdf>.

157. Copyright Act, R.S.C. 1985, c. C-42, s. 77(1).

past, the Board required applicants to file an affidavit detailing all of the steps undertaken to locate the copyright owner, in addition to other supporting documents.¹⁵⁸ Even though these requirements have been relaxed, the Board now requires much more than a declaration. According to De Beer & Bouchard:

The Board generally expects an applicant to have consulted most of the repertoires of copyright licensing agencies and collective societies, as well as national libraries' indices, copyright offices' registration records, publishing houses and corporate records. Comments made to the US Copyright Office indicated that users sometimes search on the internet, in old phone books and through death certificates and estate records. An applicant will be required to extend the search beyond Canadian borders if it is probable that the owner of the copyright may be located abroad.¹⁵⁹

Since the Act itself provides very little guidance on how the regime should work, the way the Board perceives its role is paramount. The Board emphasizes that the Act uses the terms “may issue a licence,” and not “shall issue,” giving it wide discretion.¹⁶⁰ According to the Board, it “is not under a strict obligation to issue a licence even if all the conditions have been met.”¹⁶¹ Apparently, the Board does not consider itself as a facilitator, but as a custodian, “stepping into the shoes of the absent copyright owner.”¹⁶² This means that the Board does not limit its discretion to preventing possible and foreseeable prejudices to the unrepresented party, as courts in *ex parte* proceedings normally do, but instead, the Board takes upon itself to represent the interests of the unlocatable owner.¹⁶³ This approach inevitably complicates the mechanism. Although the Board may not usually conduct a formal hearing, it may deliberate, and it may issue a reasoned decision. In addition, the Board may refuse to grant a license not because the owner was actually locatable, but because the Board concluded that the use does not actually infringe any exclusive right of the owner and therefore no license is necessary and issuing one would be *ultra vires* the power of the Board.¹⁶⁴

158. *See* De Beer & Bouchard, *supra* note 145, at 220–31 (describing the various requirements).

159. *Id.* at 228.

160. William Vancise J., COPYRIGHT BOARD OF CANADA, SPEECH GIVEN BY THE HONOURABLE JUSTICE WILLIAM J VANCISE, CHAIRMAN OF THE COPYRIGHT BOARD OF CANADA 6 (2007), *available at* <http://www.cb-cda.gc.ca/about-apos/speeches-discours/20070815.pdf>.

161. *Id.*

162. *Id.*

163. De Beer & Bouchard, *supra* note 145, at 230.

164. *Id.* at 223.

Sometimes, the Board members may disagree with each other, and a written decision with a majority and dissenting opinions would issue.¹⁶⁵ All of this is costly.

The record of decisions made under section 77 suggests that in practice the mechanism is nothing but an esoteric phenomenon. De Beer & Bouchard report that during the period between 1990 (after section 77 came into force) and 2008–2009¹⁶⁶ the Board opened 441 files pertaining roughly to 12,640 different works.¹⁶⁷ Out of these applications, 52.2% resulted in a decision (by the end of 2008, the Board granted 230 licenses and denied 5 applications), in 22.2% of the cases the owner was found, 16.3% of the applications were withdrawn, while 8.6% were not followed up.¹⁶⁸ On average, then, the Board granted at least 12 licenses per year. The total license fees set by the Board during the first 18 years of the regime was just under C\$70,000 (Canadian dollars).¹⁶⁹ The figure excludes some types of licenses, estimated to have generated royalties not exceeding a few thousand Canadian dollars.¹⁷⁰ Assuming that the total figure is C\$75,000 (or C\$4,167 annual average), the average license fee for each license was C\$326. The average is skewed upwards, because “65% of applicants sought to use only 1 work, 24% applied to use between 2 and 10 works and 7% applied to use between 11 and 100. Some applicants have sought licenses for scores of works at the same time[, while] [i]n a handful of cases, applications were made for a license or licenses covering thousands of works.”¹⁷¹

It is impossible, of course, to know how many orphan works were copied in Canada without obtaining a license from the Board (either because users are ignorant of the mechanism, or prefer risking liability over incurring the cost of applying and complying with a license), or how many uses were forgone despite the ability to obtain such a license (either because users are ignorant of the mechanism, or because they prefer foregoing the use over incurring the cost of applying and complying with a license). Yet the record

165. COPYRIGHT BOARD CANADA, ‘RE BREAKTHROUGH FILMS & TELEVISION’ 2004-UO/II-33 (Mar. 6, 2006), <http://www.cb-cda.gc.ca/unlocatable-introuvables/licences/156r-b.pdf>.

166. The report is not entirely clear about the exact period. The authors state: “The cut-off date for analysis is the end of the 2008 calendar year, although some 2009 data has been included where available and appropriate.” De Beer & Bouchard, *supra* note 145, at 242.

167. *Id.* at 242. A significant amount of these applications pertained to copying of architectural plans in municipalities’ archives (needed, for example, for building permits). In 2007, the Board concluded that no license is needed to copy those. *Id.* at 242 n.127.

168. *Id.* at 243–44.

169. *Id.* at 251.

170. *Id.*

171. *Id.* at 242–43.

so far makes it very difficult to call the section 77 mechanism a success story. Common sense strongly suggests that an average of twelve licenses per year represents a small droplet in a bucket of the actual or potential uses of orphan works in an economy of the size of Canada. The mechanism presumably has been beneficial for the applicants who obtained licenses, yet one cannot help but wonder whether the mechanism is worth preserving. The costs of maintaining the regime (for the applicants and for Canadian taxpayers) likely exceed the amount of license fees that it has generated, and even the cost of applying and processing a license likely exceeds the average license fee. It could be suggested that streamlining the procedure and increasing awareness would lead to wider utilization of the mechanism, but the very low adoption rate likely reflects a more endemic problem. It is possible that, almost by definition, the cost of the mechanism would be higher than the license fees that it would generate, no matter how streamlined the procedure is and regardless of the rate of the license fees.¹⁷² That is because the mechanism ignores the root cause of the orphan works problem: the fact that the cost of maintaining themselves locatable exceeds the license fees that owners expect to earn. If it is inefficient for owners to administer their rights under these conditions, it is unclear why the Board would be able to do that at a lower cost.

If the Canadian orphan works regime has not and cannot really provide any real solution to the orphan works problem, one cannot help but wonder what was the point in adopting and what is the point in maintaining it. Possibly, the answer is that rather than solving the orphan works problem, the purpose of the Canadian mechanism is to enshrine the dogma. Although, collaterally, in some cases, with the help of the Board, applicants were eventually able to locate the owners,¹⁷³ this is not the goal of the mechanism. Nor is there any evidence that where licenses were granted the money was

172. The license fees set by the Board are probably excessive. First, according to De Beer & Bouchard, “[t]he Board often asks collective societies for up-to-date information on the price they charge to licence their own repertoire for proposed uses of particular types of works. In markets where collective administration does not exist, it is often possible to determine a generally recognised market practice.” *Id.* at 235. Since the prices that are set by collective societies often reflect their monopolistic position and absence of competition, using their prices are a poor proxy for the license fees that an owner and user would agree on in an arms-length competitive negotiations. Second, using prices set for non-orphan works as a proxy for the price of orphan works is misguided because, as noted above, the reason why the work is orphan is that the owner does not expect to earn license fees that exceed the cost of licensing. That is, rather than relying on the revealed preferences of other copyright owners, the Board should have relied on the preferences of the copyright owner, who effectively revealed that she does not expect to earn any meaningful license fees.

173. *Id.* at 243.

actually collected by the owners. The mechanism does not solve the orphan works problem in the sense of matching between users and owners. Instead, the mechanism offers users that are either extremely risk averse or are devote believers in the copyright dogma an ability to obtain indulgence, while entrenching the notion that using work is a sin, and any benefit derived from it is sinful.

The clearest indication that the section 77 mechanism is—or at least has become—an instrument for the sale of indulgences is the Board's controversial practice of ordering licensees to pay license fees to a collecting society.¹⁷⁴ Under section 77(3):

The owner of a copyright may, not later than five years after the expiration of a licence issued pursuant to subsection (1) in respect of the copyright, collect the royalties fixed in the license or, in default of their payment, commence an action to recover them in a court of competent jurisdiction.¹⁷⁵

The implication is that a user who successfully applies for a license becomes immune from an infringement action, and is only liable to pay the license fees determined by the Board if, and only if, the copyright owner emerges within five years after the expiry of the license. If the owner never shows up, as might be expected (or shows up later than five years after the expiry of the license), the user will not be required to pay anything, will be liable for nothing, and will be under no obligation to share the benefit arising from using the work. Nor does the user have to hold the license fees in trust or pay an escrow. The owner, in return, benefits from an extended limitation period, and from an ability to recover the license fees in an expeditious process (instead of suing for infringement) should the user refuse to pay the prescribed fees.

Despite what seems to be clear direction from Parliament that the payment is completely contingent upon the owner's emergence within five years, the Board decided to order applicants to pay copyright collectives on a non-contingent basis, immediately upon the issuance of a license.¹⁷⁶ The Board does not even require the copyright collective to hold the license fees in trust (it did so in the past), but allows the collective to use the money as it sees fit, as long as the collective undertakes to compensate the owner if it becomes necessary.¹⁷⁷

174. *Id.* at 236.

175. Copyright Act, R.S.C. 1985, c. C-42, s. 77(1) (Can.).

176. *See* De Beer & Bouchard, *supra* note 145, at 236.

177. *Id.*

What justifies such a radical shift from the mechanism contemplated by Parliament (or, perhaps more precisely, such a subversion of that mechanism)? De Beer & Bouchard mention the following rationales: (1) anecdotal evidence that involving collective societies in the process increases the likelihood that royalties will eventually reach the true copyright owner; (2) the possibility that it might be easier for copyright owners to collect owing amounts from a collective than from the user who might disappear; and (3) the potential to help collective societies defray some of the costs that they incur in cooperating with the Board on section 77 applications.¹⁷⁸

These arguments are not convincing. While it seems likely that collective societies have information that can facilitate locating copyright owners, it does not follow that the collective should get paid when people use works of owners who never authorized the collective to license on their behalf. By definition, an orphan work cannot be part of the collective society's repertoire: if it were, it would not be an orphan work. At most, the fact that collecting societies might possess information that might facilitate locating owners may justify making the data that collectives have available to the public.¹⁷⁹ Similarly, the fact that sometimes the owner might find it easier to collect the money from a collective than from the user is not self-evident, and even if it were, does not justify departure from what Parliament had contemplated. In fact, the fact that users who apply for a section 77 license may practically not pay anything is not a bug of the mechanism, but a feature, because it encourages users to apply a license. The Board's decision to require immediate upfront payment to a collective society punishes the very few applicants who actually go into the trouble and cost of applying for a section 77 license.¹⁸⁰

Bouchard, who is also the Copyright Board's General Counsel, in a recent paper offered additional, and more forthcoming, justifications for the

178. *Id.* at 238.

179. In fact, under sections 67 and 70.13 of the Canadian Copyright Act, collective societies must answer within a reasonable time all reasonable requests from the public for information about their repertoire of works. Copyright Act, R.S.C. 1985, c. C-42, s. 67, 70.13 (Can.).

180. Theoretically, if there is no statutory basis for the Board's decision to order upfront payment to a collecting society, such an order could be set aside by the Federal Court of Appeal following an application for judicial review, *see* Federal Courts Act, R.S.C. 1985, c. F-7, s. 28(1)(j) (Can.). Realistically, such an application for judicial review seems very unlikely, as long as the license fees that the Board orders to pay are not excessively high, because a user who is willing to incur the cost of judicial review is probably better off not seeking a license in the first place and risk incurring the cost of litigation should the copyright owner ever emerge.

decision to order applicants pay collective societies.¹⁸¹ These justifications reflect the view of the user as a sinner, and betray the Board's allegiance to the copyright dogma. Bouchard explains that while the Board acknowledges that non-contingent royalties payable to a collective society are controversial, the Board's position is that the user must generally be required to pay because the Board "does not believe that it should be in the business of issuing free insurance policies against prosecutions for violation of copyright."¹⁸² He explains that the Board compares the user to a defendant in a class action who, after being found liable, may be disgorged of the surplus unclaimed by the class,¹⁸³ a comparison that reveals how the Board treats the user as a wrongdoer whose activities should be discouraged, not facilitated. Finally, he notes that the Board relies on the concept of *cy pres* in trust law, whereby a court may order that property be used for a purpose that is as close as possible to the use originally intended if it becomes impossible to carry out the original purpose.¹⁸⁴ Even though Parliament clearly contemplated that if the owner does not emerge the user can keep the surplus, the Board views disgorging the user of this surplus as a closer purpose because "given the choice, the unlocatable copyright owner would prefer that the royalties be paid to a group that represents interests similar to those of the owner than to see the user take advantage of the owner's copyright for free,"¹⁸⁵ and "when a protected use of a protected work is contemplated, the payment of royalties should be the norm, not the opposite."¹⁸⁶

These statements present some of the clearest articulations of the copyright dogma. It is far from obvious that the average copyright owner would rather funnel such fees to benefit other authors¹⁸⁷ than be retained by the user. In fact, one could assume that many owners might actually be quite happy, if not thankful, if someone showed interest in their works after they ceased to be commercially viable and exhumed them from oblivion. The fact that the Board unequivocally assumes that the average copyright owner

181. Mario Bouchard, *The Canadian Unlocatable Copyright Owners Regime*, in THE COPYRIGHT BOARD OF CANADA: BRIDGING LAW AND ECONOMICS FOR TWENTY YEARS 137, 153 (2011).

182. *Id.* at 153.

183. *Id.*

184. *Id.*

185. *Id.*

186. *Id.* at 154.

187. Also note the subtle reference to the beneficiaries as "authors" rather than "owners" even though the main beneficiaries of fees collected by collecting societies, other than the societies themselves, tend to be publishers, not authors.

would a priori prefer another owner over the user reflects the underlying view that any unpaid use is inherently sinful. Effectively, by requiring users who apply for a license to pay a collective society, the Board has torn apart the essential feature of copyright as a set of bilateral relationships between owners and users, and created, without any statutory mandate and contrary to the mechanism contemplated by Parliament, an extended collective licensing mechanism. The merit of that mechanism is discussed in the next Section.

C. EXTENDED COLLECTIVE LICENSING

Extended Collective Licensing (“ECL”) regimes have existed in some of the Nordic countries since the 1960s.¹⁸⁸ In the aftermath of the rejection of the Amended Settlement Agreement (“ASA”) in the Google litigation,¹⁸⁹ they have gained increasing popularity as a possible solution to the orphan works problem.¹⁹⁰ Under an ECL regime, a collecting society that represents a substantial number of rights holders may grant licenses that extend not only to copyrighted works owned by the society’s members, but also, with respect to the same set of uses, to works owned by non-members. The society is typically authorized to act on behalf of foreign as well as domestic rights holders.¹⁹¹ In some cases copyright owners may opt out, and therefore will not be bound by the agreement entered by the collecting society, but in some other cases they cannot.¹⁹²

ECL regimes were not designed to solve an orphan works problem, but ECL is considered by some to be an attractive solution for the orphan works problem because it solves the problem in one fell swoop.¹⁹³ Indeed, the orphan works problem is simply assumed away, as the ECL, by definition, includes the orphan works in the collecting society’s repertoire.

Moreover, ECL could be seen as a logical extension of the rationale for collective administration generally. The standard justification for collective administration of copyrights is the reduction of the transaction costs of licensing and enforcing the copyrights in situations when administering these

188. Thomas Riis & Jens Schovsbo, *Extended Collective Licenses and the Nordic Experience: It’s a Hybrid but is it a Volvo or a Lemon?*, 33 COLUM. J.L. & ARTS 471 (2009).

189. *Authors Guild v. Google, Inc.*, 770 F. Supp. 2d 666, 686 (S.D.N.Y. 2011).

190. See Samuelson, *supra* note 6, at 705–15; see also Robert Darnton, *Jefferson’s Taper: A National Digital Library*, N.Y. REV. BOOKS (Nov. 24, 2011), available at <http://www.nybooks.com/articles/archives/2011/nov/24/jeffersons-taper-national-digital-library/>.

191. Riis & Schovsbo, *supra* note 188, at 475.

192. *Id.* at 476.

193. See sources cited *supra* note 190.

rights in a competitive market is assumed to be prohibitively costly, or at least highly inefficient.¹⁹⁴ That is why Richard Posner calls them “benign cartels.”¹⁹⁵ Assuming that there are indeed situations where competitive administration of *commercially* available works is indeed prohibitively costly and collective administration is desirable,¹⁹⁶ then it seems only to make sense to include orphan works in the collective’s repertoire because the cost of those works through the market is prohibitively costly a fortiori.

Beyond this superficial appeal, however, it is doubtful whether ECL is an appropriate solution to the orphan works problem. Note that the logic of the previous paragraph does not work in reverse. If prohibitively high transaction costs in a particular industry justifies collective administration of copyrights of works whose owners are known, there is merit in extending collective administration to orphan works, because the transaction costs involved in licensing those are, by definition, even higher. However, it does not follow that when the market fails in a subset of an industry (e.g., orphan works) it makes sense to solve this failure by extending that solution to all other parts of that industry where the market works just fine. In other words, when there is, or could be, a reasonably well-functioning competitive market for non-orphans (which are, presumably, the more commercially valuable ones), it does not make sense to replace that competitive market with a monopolized structure in order to solve the problem of the orphans (which are presumably the least commercially valuable ones).

Consider the ASA, which, had it been approved, would, in effect, have created a *sui generis* ECL.¹⁹⁷ The forward-looking element of the ASA consisted of two separate parts: first, it proposed to create The Book Rights Registry (“Registry”), which would be, effectively, a new collecting society for works whose copyright owners are known and locatable (“locatables”). In addition, the ASA offered a solution to the orphan books problem (books whose copyright owners are unknown or unlocatable). The first part involved a problematic solution to a problem that does not exist. The second part, however, involved a solution, albeit flawed, to a real problem. Both aspects

194. How often this happens is a separate question. See generally Ariel Katz, *Copyright Collectives: Good Solution But for Which Problem?*, in *WORKING WITHIN THE BOUNDARIES OF INTELLECTUAL PROPERTY: INNOVATION POLICY FOR THE KNOWLEDGE SOCIETY* (Harry First, Rochelle C. Dreyfuss, & Diane L. Zimmerman eds., 2009).

195. RICHARD A. POSNER, *ANTITRUST LAW* 30 (2001).

196. But see Katz, *supra* note 194 (questioning the prevalence of situations that justify collective administration).

197. See Samuelson, *supra* note 6, at 706–08 (discussing the similarities and differences between the ASA and ECL).

were subject to an opt-out rule—namely, they would not apply to copyright owners who actively opt-out of the settlement.¹⁹⁸

The first aspect of the ASA, dealing with the sale of locatables, is difficult to justify. The Registry would have been essentially a new collecting society with the power to collectively set the prices and terms for the sale of digital books. However, it is difficult to see why collective administration is desirable in a market for digital books. The perceived market failures that might justify collective administration in some narrow circumstances¹⁹⁹ do not generally apply to a project like Google Books or similar projects. First, there are no prohibitive transaction costs preventing individual copyright owners and Google from transacting over the terms of commercializing their works.²⁰⁰ Indeed, even the smallest copyright owner could locate and contact Google if he wants to (and if he does not know how to find Google, he can google Google). Second, it is very easy for a copyright owner to monitor whether Google has used the work without authorization (he can google the work and immediately find out; he can even create a “Google Alert” and be notified automatically).²⁰¹ If striking deals with locatable copyright owners works for Amazon or Apple with respect to books and sound recordings, then it should work for Google, too.²⁰² There is no market failure that may justify collective administration (and the associated reduction in competition) in the case of in-print works (or even out-of-print works whose owners are known and locatable). Therefore, the ASA offered an anti-competitive solution to a problem that does not exist.²⁰³ The sale of digital copies of such books should be subject to copyright law’s ordinary opt-in (i.e., permission-first) rule.²⁰⁴

198. *ASA*, *supra* note 2 (Attachment I, Notice of Class Action Settlement, 25).

199. *See Katz*, *supra* note 194.

200. *See Randal C. Picker*, *The Google Books Settlement: A New Orphan Works Monopoly*, 5(3) J. COMPETITION L. ECON. 383, 403 (2009).

201. *Google Alerts*, GOOGLE, <http://www.google.com/alerts> (last visited Sept. 22, 2012).

202. On October 4, 2012, Google and the plaintiff publishers in the Google Books litigation reached a settlement that allows Google to continue doing what it had done before and allows the publishers to choose whether Google will display larger parts of their books and sell copies of them. *See Claire Cain Miller*, *Google Deal Gives Publishers a Choice: Digitize or Not*, N.Y. TIMES, Oct. 5, 2012, at B7. The settlement, and the fact Google has offered such “partnership” agreements all along, confirm that there was no market failure that would hinder such transactions.

203. *Picker*, *supra* note 200.

204. The *sale* of digital books should not be confused with the original function of Google Books: to serve as a searchable index of books. As Sag notes, Google’s scanning and caching of books for the purpose of providing a searchable index should not constitute copyright infringement at all. *See Sag*, *supra* note 48.

The case for orphan works is different, and the part of the ASA dealing with orphan works tried to solve a real market failure. Had it been approved, the ASA would have permitted Google to sell copies of orphan works until their owners showed up and demanded otherwise.²⁰⁵ Meanwhile, Google would transfer 63% of its earnings from the sale of such books not to the copyright owners (because they are unlocatable), but to the Registry, which, with the aid of the Unclaimed Works Fiduciary, would attempt to locate these owners and distribute the money to them (after deducting its expenses).²⁰⁶ If those owners could not be found, the money would eventually be remitted to literary based charities. Effectively, the ASA contemplated an ECL regime.

One of the main objections to this aspect of the ASA was that it would effectively grant Google a monopoly over the commercial use of orphan works.²⁰⁷ This was a serious concern, but if the alternative to monopoly (one seller) is “nullopoly” (no seller), then a monopoly is probably the better option.²⁰⁸ However, presenting the question as a choice between monopoly and nullopoly is false: it assumes that the only realistic alternative to an orphan works nullopoly is a monopoly,²⁰⁹ and it further assumes that monopoly in the form of ECL solves the orphan works problem. But an orphan works monopoly is not the only alternative to a nullopoly because the nullopoly can be averted by other means, such as this Article’s remedy-tweaking solution. And moreover, ECL does not really solve the problem, it only pretends to. ECL relieves some pressure off the absurd consequences of copyright dogma but a solution to market failure it is not, as the next Section explains.

205. *ASA*, *supra* note 2.

206. *Id.*

207. *See, e.g.*, Statement of Interest of the United States of America Regarding Proposed Amended Settlement Agreement at 21, *Authors Guild v. Google, Inc.*, 770 F. Supp. 2d 666 (S.D.N.Y. 2011) (No. 05 Civ. 8136), *available at* http://thepublicindex.org/docs/amended_settlement/usa.pdf (“A core issue that arises out of the parties’ effort to resolve this matter is the ability of Google, and no other entity, to compete in a marketplace that the parties seek to create. Nothing in the ASA addresses this concern. . . . There is no serious contention that Google’s competitors are likely to obtain comparable rights independently.”); Picker, *supra* note 200, at 385–86.

208. *See* Einer Elhauge, *Why the Google Books Settlement is Procompetitive*, 2 J. LEGAL ANALYSIS 1 (2010).

209. *See* Randal C. Picker, *Antitrust and Innovation: Framing Baselines in the Google Book Search Settlement*, GCP: THE ANTITRUST CHRONICLE 5 (Oct. 2009), *available at* <http://www.law.uchicago.edu/files/file/493-rcp-google.pdf>.

D. DOES ECL EVEN SOLVE THE ORPHAN PROBLEM?

Recall that the orphan works problem stems from the inability of owners and users to transact. As discussed above, this problem has two sides: the demand side (the user) and the supply side (the owner). ECL supposedly solves the orphan works problem because it allows the user to obtain a license not from the owner, who is by definition unlocatable, but from a third party, such as a collecting society. If this option is available, then it partly solves the demand side of the problem (I say partly because the user needs to deal with a monopoly, which creates another set of problems that I address below). But requiring the user to seek permission from and pay a fee to a collecting society (or even only pay a fee) does not really solve the supply side of the problem. Instead, it only shifts the problem from one location to another: the problem of locating the unlocatables moves from the user to the collecting society. But shifting a problem does not necessarily solve it. Of course, we can pretend that it does, and as a matter of positive law, if lawmakers decide that instead of getting permission from the copyright owner a user can pay a fee to a collecting society, then the problem, as a practical matter, may be solved, but it is solved not because the solution addressed the causes of the problem but because the lawmaker decides that an act that would normally attract liability may be absolved by paying a fee to a collecting society.

Why lawmakers should adopt such a rule, however, is far from clear. Perhaps, if there were reasons to believe that the collecting society would only request the fee that the owner would have asked, and that upon payment, the money would miraculously find its way to the owner (who is, after all, unlocatable), then the collecting society would be a perfect intermediary and ECL would be a perfect solution. But realistically, it is more likely that the collecting society will not become that perfect intermediary. It is more likely that the collecting society will charge supra-competitive fees that will discourage use of orphan works, and that if any money will eventually be distributed to the copyright owner, most of the collected fees would benefit others: the collecting society's insiders, other copyright owners, or, in the case of the Fiduciary, literary-based charities. It is difficult to see how requiring users to pay a fee to someone other than the copyright owner advances the purposes of copyright. Such a payment is a windfall for the recipient, and therefore cannot provide any additional incentive to create beneficial works,²¹⁰ and from the user side, such a fee constitutes an

210. See Randal C. Picker, *Private Digital Libraries and Orphan Works*, 27 BERKELEY TECH. L.J. 1259, 1277–79 (2012).

additional cost, and therefore it would decrease, rather than enhance, access to and dissemination of orphan works.

ECL is an acceptable solution to the orphan works problem only if we no longer think about copyright in terms of the correlative rights and duties of owners and users and their private reordering in open markets, and stop asking how those correlative rights and duties can be exercised and complied with in ways that promote the public interest and minimize interference with the rights and liberties of others. Instead, ECL is an acceptable solution only if we define copyright solely by the prohibitions it imposes on users, if we view copying another person's work as inherently sinful, regard prior permission as the means to avoid and absolve that venial sin of copying, and admit that we may not be interested in avoiding sin as we are interested in allowing collecting societies to monetize it, through the sale of indulgences.

E. ECL AS INSURANCE

One defense of ECL is that it is a pragmatic solution, which allows risk-averse users to pay a fee and eliminate the risk of liability. Essentially, it could be argued, ECL is an insurance scheme, whereby in exchange for a fee, the collecting society relieves the user from her potential liability and assumes an obligation to pay the owner should he emerge. Yet, this insurance scheme is quite extraordinary for several reasons. First, if the law disapproves of the unauthorized use of a work by regarding such use as an infringement, it is questionable whether a contract under which one party undertakes to indemnify another for the risk of *deliberate* infringement is valid, because contracts that "create incentives to commit acts that society has made illegal or, at least, disapproves of" may be against public policy.²¹¹ There might be reason to rethink this rule, especially if relaxing it could help solve the orphan works problem. However, even if it is acceptable that the user can be relieved of her duty to seek permission before using a work by paying a fee to someone other than the copyright owner, it is not self evident that the collecting society should be the preferable payee. If allowing users to insure themselves against the risks arising from unauthorized use of orphan works is a good solution, then there is no reason to designate a collecting society as

211. See *Truck Ins. Exch. v. Ashland Oil, Inc.*, 951 F.2d 787, 790 (7th Cir. 1992). This should be distinguished from Errors and Omissions insurance policies that may cover the risk of non-deliberate infringement claims. See also Howard P. Knopf, *Copyright Collectivity in the Canadian Academic Community: An Alternative to the Status Quo?*, 14 I.P.J. 109, 122 (1999) (noting that the legislation was needed to allow responsible publishers to obtain libel insurance because contracts for indemnification of illegal acts, such as contracts for libel insurance, were unenforceable under the common law).

the sole insurer. It would be better to have a competitive insurance market instead.

VIII. THE POLITICAL ECONOMY OF THE COPYRIGHT DOGMA

A discussion of solutions to the orphan works problem will not be complete before addressing why has it been so challenging to find an acceptable and workable solution to this problem in the first place. The difficulty is puzzling because owners of orphan works are, by definition, absent from the debate about orphan works, and normally, when discussions about contemplated reforms do not involve those who might be directly affected by them, one could expect that reform would be easy. Specifically, one could expect that in a setting where users lobby for reform that would allow them to use orphan works, and owners of those works are absent, passing a pro-user reform (even overly pro-user) would be a breeze. Therefore, the fact that it has been difficult to find an acceptable solution, and that many of the proposed solutions involve serious impediments on using orphan works, suggest that the political economy of the orphan works problem is complicated, and that there is much at stake—not necessarily for the interests of orphan owners, but for the interests of those who speak on their behalf. This Section looks at some of these interests in order to understand who stands to win and lose from different solutions, and who has an interest in promoting the copyright dogma. This Section shows that besides some legitimate concerns, anti-competitive or opportunistic motivations drive many of the objections and the proposed solutions.

Answering the puzzle begins with the recognition that the interests of those who claim to speak on behalf of the orphans and the interests of the orphans themselves may not always be aligned. Evidently, when Amazon's counsel argued before Judge Chin that a copyright owner may sit back and do nothing, he was not defending the interests of the orphans or of any other copyright owner for that matter;²¹² he must have been more concerned about how Google Books might affect his client's online retail business. But even the interests of known copyright owners and their organizations may not be fully aligned with the interests of the owners of orphan works. Indeed, there are several ways in which promoting the copyright dogma benefits the interests of the known, commercial, or organized copyright owners' groups in ways that do not necessarily advance the interests of orphan owners. Some of their concerns are legitimate, while others are less so.

212. *See* Authors Guild v. Google, Inc., 770 F. Supp. 2d 666, 681 (S.D.N.Y. 2011).

Known owners may be legitimately concerned about false positives. Many proposed solutions depend on defining conditions under which a work may be designated as “orphan.” This creates a risk that courts would treat some works as orphans even when their owners are locatable.²¹³ Although this concern is legitimate, eliminating the risk may increase the likelihood of false negatives, namely, works that should have been treated as orphans but are not. Indeed, this is what happened to the Copyright Office proposal from its 2006 Report, to the 2008 Bill, as more stringent requirements on users were added before a work could be treated as a “orphan.” While these changes might have reduced the risk of false positives, they increased the risk of false negatives, and therefore might excessively deter the use of orphan works.²¹⁴

Other objections to simple and workable solutions may be motivated by less legitimate concerns. Some copyright owners have nothing to gain from any change in the status quo, and therefore perceive any deviation from it as potentially harmful. Recall that relaxing the permission-first rule by recognizing a duty to make oneself locatable can benefit future users, the public in general, as well as present owners (by encouraging reuse and revival even if owners let their works become orphans). But relaxing the rule does not benefit all owners to the same degree, because owners of non-orphan works are those who, by definition, believe that the present value future uses is larger than the cost of making themselves locatable. Therefore, these owners are unlikely to benefit from change to the status quo of a strict permission-first rule, and might perceive any deviation from it as potentially harmful.

An additional motivation for objection could be termed as anti-competitive, while the last one can be called opportunistic. A strict permission-first fosters a “clearance culture”²¹⁵ that can function as a barrier to entry and protects large copyright owners from competition in several markets. Entities who own or administer large portfolios of works, such as big publishers, or the music and film “majors,” have comparative advantage in making themselves locatable and maintaining ownership information

213. See Edward Hasbrouck, *Facts and Fallacies of “Orphan Works,”* THE PRACTICAL NOMAD BLOG (Apr. 4, 2012, 12:02 PM), <http://hasbrouck.org/blog/archives/001996.html>.

214. See *supra* Part VII.

215. PATRICIA AUFDERHEIDE & PETER JASZI, UNTOLD STORIES: CREATIVE CONSEQUENCES OF THE RIGHTS CLEARANCE CULTURE FOR DOCUMENTARY FILMMAKERS 22 (2004), available at <http://www.acsil.org/resources/rights-clearances-1/nps240.tmp.pdf> (defining “clearance culture” in the context of documentary filmmaking as “the shared set of expectations that all rights must always be cleared—both imposed upon filmmakers and imposed by them on themselves and colleagues”).

compared to small owners. A strict permission-first rule makes it more difficult and risky to use works other than those that these entities can license and therefore benefits them by making the licensing markets less competitive.

Many of these entities are also producers of content, and the clearance culture that the dogma fosters creates a barrier to entry in the markets for their produced content. While they may occasionally benefit from a less stringent a permission-first rule when reusing others' works, large producers might still find the copyright dogma worth defending because they can easily thrive in a clearance culture, whereas smaller producers and new entrants may find this culture very difficult to navigate.²¹⁶

Another anti-competitive motivation is the concern that under some conditions orphan works could serve as substitutes for non-orphans. Consequently, reforms that may increase the supply of usable orphan works will erode the market power of the non-orphans. Indeed, objections of this kind were made explicit in debates in France. In its submission with respect to solutions to the orphan works problem, the French Centre Français d'exploitation du droit de Copie ("CFC"), a collecting society that licenses books and press reproduction rights, demanded that the use of orphan works would be permitted only if compensation is made to a collecting society, because (a) otherwise, orphan works freely accessible online would create "unfair competition" with works that are commercially available (presumably, the works of CFC members); and (b) permitting orphan works to be freely accessible online would devalue the content of commercially available works and encourage users to demand free access to all works.²¹⁷ In other words,

216. Jeremy Williams, *Who is concerned about broader access to orphans and why?*, Berkeley Orphan Works Symposium, audio available at http://media.law.berkeley.edu/qtmedia/BCLT/bclt_20120412-symposium/day1/Williams.m4a (Jeremy Williams from Warner Brothers Entertainment, Inc. described at the symposium how unlike small producers, Warner's large legal department is capable of clearing rights).

217. Bernard Lang, *L'exploitation des oeuvres orphelines dans les secteurs de l'écrit et de l'image fixe* [Annex to the CSPLA Report], 19 (Mar. 17, 2008), <http://www.datcha.net/orphan/oeuvres-orphelines-BLang.pdf> (citing *Les oeuvres Orphelines dans le Secteur de l'Écrit*, Groupe de travail CFC, note d'étape, 2, 2 octobre 2007, available at <http://bat8.inria.fr/~lang/orphan/documents/france/CFC-NOTE-D-ETAPE-2007.10.02.pdf>). Interestingly, in the recent eBooks antitrust litigation, the defendant publishers made similar arguments. One of their justifications for agreeing with Apple on an "agency" pricing model was the concern that the low prices that Amazon charged for e-books under the previous "retail" model would devalue hardcover books, reduce consumers' willingness to pay higher prices for hardcover books and force publisher to lower the retail price of those books. See Jonathan M. Jacobson & Kimberley A. Piro, *Storm over E-books Waters*, in *COMPETITION LAW JOURNAL: CONCURRENCES* 14, 14–15 (2012), available at <http://ssrn.com/abstract=2140778>.

legally usable orphan works will increase the supply of usable works, which will force down the prices set by CFC or its members.

The third motivation can be termed as merely opportunistic. The motivation here is not a concern about harm (whether cognizable or not) but about seizing an opportunity to profit: realizing that there is value in the use of orphan works, the objection is aimed at ensuring that this value would be captured by the objector and not by others.

The entities that have the clearest interest in promoting the copyright dogma are collecting societies, who can bring the anti-competitive and the opportunistic motivations to perfection. Unlike their members, some of which may sometimes find themselves on the user side, collecting societies neither produce content nor distribute it. Their business consists merely of licensing content already produced and delivered by others, and they are driven by a single objective: to maximize the license fees that they collect. For collecting societies any market failure, whether real or perceived, presents a business opportunity in the form of a new area of activity that they will be prepared to collect money for.²¹⁸ Consequently, they view any reform that would facilitate transactions in a competitive market as a lost opportunity to collectively license this activity, and any proposal to exempt some activities that they are happy to license as an existential threat.²¹⁹ As organizations driven by a single purpose, they do not need to trade off the inefficiencies that the rules that they promote create, and even when those

218. David Vaver, *Commentary*, in PROCEEDINGS OF THE COLLOQUIUM ON THE COLLECTIVE ADMINISTRATION OF COPYRIGHT 161, 166–68 (1994) (on file with author).

219. For example, Access Copyright, a Canadian collecting society representing publishers and some authors has lobbied forcefully against any expansion of Canada's fair dealing provision, and in particular, against the explicit inclusion of "education" as a recognized purpose. See ACCESS COPYRIGHT, ANNUAL REPORT 2011 7 (2012), http://www.accesscopyright.ca/media/22585/2011_annual_report.pdf. According to a 2007 report, commissioned by Access Copyright, and written by Prof. Martin Friedland, more than 75% of Access Copyright's income came from licensing the educational sector (K-12 and post-secondary), in contrast to the U.S.-based Copyright Clearance Center, whose income from the education sector amounted to only 12% of its income and has not operated in the K-12 area at all. See MARTIN FRIEDLAND L., REPORT TO ACCESS COPYRIGHT ON DISTRIBUTION OF ROYALTIES 19 (2007), *available at* http://www.accesscopyright.ca/media/8359/access_copyright_report_-_february_15_2007.pdf. Remarkably, despite the fact that the population in Canada is about one-tenth of the U.S. population, Access Copyright collected more money from the education sector than the CCC did. According to Friedland, in 2006 CCC collected approximately \$19 million from the education sector. *Id.* at 24. Access Copyright total revenue in 2006 was C\$34.2 million, meaning that it collected approximately C\$27 million from the education sector (or \$23.6 million). See ACCESS COPYRIGHT, ANNUAL REPORT 2007 10 (on file with the author). It is, therefore, clear why Access Copyright lobbied forcefully to prevent the explicit recognition of "education" in the fair dealing provision.

inefficiencies affect their members, the additional profit arising from collecting societies' ability to collect monopolistic fees for the use of their members' works may compensate the members for the losses that those inefficiencies create.

Naturally, collecting societies are eager to become the guardians of orphan works. The ability to charge users a fee for the use of orphan works through an ECL mechanism allows them: (a) to control the supply of orphan works and ensure that they do not compete with the works of their members (the anti-competitive motivation); and (b) to profit from, and appropriate the value of, the use of the works of others (the opportunistic motivation). While the benefits of ECL to the collecting society and its members may be obvious, it is less obvious that installing a monopoly of non-orphans as a custodian and gatekeeper of orphans benefits users and the public at large. The problem is that the stakes for the collecting societies (and their members) and users are asymmetric. Under competitive solutions, users can only expect to earn normal returns, whereas under ECL the collecting society and its members can earn supra-competitive returns. The copyright owners, because they can gain more, can rationally spend more resources on lobbying for a monopolistic solution than users would spend on lobbying for competitive solutions.²²⁰ Moreover, since the owners of orphan works are by definition not members of collecting societies, it far from clear why a collecting society would be entrusted with their guardianship. Therefore, to promote themselves as the only acceptable solution, collecting societies (or owners seeking to establish new ones) need to promote the notions that permission-first is the unquestionable tenet of copyright, that the orphan works problem is an insurmountable one, and that they are the only ones who can come to the rescue. In other words, to sell ECL, its proponents need to frame the problem in terms that define ECL as the only possible solution.

Therefore, the move to ECL requires not only promoting permission-first as fundamental tenet of copyright, but it also requires some doctrinal innovation that transforms copyright law into a copyright dogma. What distinguishes the copyright dogma from copyright law is that the dogma abandons the notion of copyright as a set of bilateral and correlative rights and duties of owners and users and defines copyright solely by the prohibitions it imposes on users. The dogma removes copyright law from the

220. *Cf.* Patry & Posner, *supra* note 32, at 1645. Patry and Posner identify these asymmetric stakes as a general problem in copyright. Collecting societies aggravate the problem because they generate industry-wide monopoly rents above the rents that an individual copyright owner could get by virtue of the copyright.

realm of private law and thus removes the constraints that the bipolar relationships between owners and users impose on “the very scope and structure of both the action and the entitlement.”²²¹ As Balganesch explains:

The bipolarity of private law entitlements . . . ensures that a plaintiff will seek relief only against a specific defendant (rather than the state or society), that a defendant will be ordered to compensate a specific plaintiff (rather than the collective), and perhaps most importantly, that the relief sought and recovery obtained will be determined entirely by the parties’ actions rather than by any external determinants.²²²

The dogma serves collecting societies and their members because it transforms the wrong of copyright infringement from copying that *harms* the copyright owner to copying that is harmful per se; from copying without the permission of *the* copyright owner, to copying without permission of *a* copyright owner; and then from copying without paying *the* owner, to copying without paying *a* copyright owner, or better yet, copying without paying an organization of copyright owners. This kind of dogmatic innovation allows collecting societies to transform copyright law into an apparatus for monetizing sin, which brings us back to the sale of indulgences by the Catholic Church.

In the provocative book *Sacred Trust: The Medieval Church as an Economic Firm*, Robert Ekelund and his coauthors analyze the medieval Catholic Church as a corporation that promoted and offered a “product” that many consumers felt they had to have: the salvation of their immortal souls.²²³ The Catholic Church was not the only religion offering this product, and one of its methods to monopolize this market was doctrinal innovation. The invention of the purgatory—a temporary state between heaven and hell—was a central element in this doctrinal innovation.²²⁴ Ekelund et al. describe the doctrine as follows:

The Church held that sins had to be paid for either in this world or the next. Payment was proportioned to guilt. Grievous offenses (i.e., mortal sins) condemned one to hell if the offender died unrepentant. Purgatory was of no consequence in such cases. However, sinners who repented and received absolution from a priest retained the opportunity to enter heaven, provided they did

221. Balganesch, *supra* note 25, at 1667.

222. *Id.* at 1667–68.

223. ROBERT B. EKELUND ET AL., *SACRED TRUST: THE MEDIEVAL CHURCH AS AN ECONOMIC FIRM* 152 (1996).

224. *Id.* at 153.

penance. Before purgatory came to be official doctrine, sins could be atoned on earth only by good works, that is, payment rendered directly within one's earthly lifetime. The invention of purgatory essentially introduced a means of "deferred payment," which not only allowed atonement to be postponed beyond this life, but also allowed third parties to make payments on behalf of the deceased.²²⁵

Initially, sins could be atoned only by good works, but over time, further innovation occurred, whereby monetary payment, obtained by the sale of indulgences, was accepted in lieu of good works.²²⁶ Assuming that with some exceptions most potential believers have some affinity to commit venial sins, these innovations reduced the price of sin, which made the product of the Catholic Church more attractive than that of its competitors.²²⁷ Ekelund et al. ask, "[B]ut why would the Church introduce a doctrinal innovation that encouraged sin?", and offer an economic answer: "The Church wanted to increase its membership and prevent existing members from defecting to alternative religions that offered assurances of salvation at a lower price. In other words, at least in part, purgatory is a market response to doctrinal competition."²²⁸

Essentially, through doctrinal innovation the Church was able to monetize sin. Not only it was able to outcompete other religions who charged a higher price for sin, the business of monetizing sin through the sale of indulgences proved to be highly lucrative. The doctrine further ensured that the authority over the sale of indulgences, and the opportunity to profit, would be concentrated in the hands of the Vatican.²²⁹

Discussing the merit of medieval Catholic dogma is certainly beyond the scope of this Article, but the elements of monetizing sin through the sale of indulgences is common to ECL, the Canadian regime, and similar proposals. Promoting the idea that reusing orphan works is inherently sinful, and that to save their souls from the copyright hell of strict liability users should purchase indulgences from the Copyright Church, may enrich the institutions of the Copyright Church. Meanwhile, it preserves the flawed notion that any use of a work without permission is wrongful, hinders the emergence of competing visions of copyright law, and suppresses the resurrection of the orphans.

225. *Id.* at 154–55.

226. *Id.* at 155, 157.

227. *Id.* at 156.

228. *Id.* at 157.

229. *Id.* at 157–58.

IX. CONCLUSION

In general, copyright requires users to seek permission before the copy protected works. Generally, the permission-first rule makes sense because it supports copyright's market orientation, but when permission-first leads to a market failure copyright law should, and often does, deviate from it.

The orphan works problem is just one instance of a case in which adhering to a strict permission-first rule leads to a market failure, but solving this market failure does not require any grand solution. No rewriting of copyright's basic tenets is necessary, and there is no need to create new monopolies, bureaucracies, or a combination of both. Solving the orphan works problem requires recognizing its causes. The problem arises not only because users find it prohibitively costly to locate owners, but because under a strict permission-first rule owners, who do not internalize the full social cost of forgone uses, face suboptimal incentives to maintain themselves locatable. Yet copyright owners are generally the least-cost avoiders of this problem and solving the problem requires adjusting the users' duty to seek permission with the owners' duty to maintain ownership information available. This is not a radical proposal: it is elementary. Imposing a duty to reduce avoidable costs on the party who is best suited to avoid them is prevalent throughout the law. If lawmakers feel impeded from applying this principle in copyright, it is only because they have been blinded by the indoctrination of the copyright dogma.

If owners do not presently face sufficient incentives to maintain themselves locatable, limiting the full array of remedies for failure to remain locatable will encourage them to remain locatable. Remedy tweaking will shift some of the risk of using orphan works from users—who currently bear it entirely—to owners. Creating incentives for keeping ownership information up to date will reduce the scope of the orphan works problem, without resorting to solutions that undermine copyright law's market orientation, and require costly procedures and unnecessary bureaucracies.

The low cost of making themselves available is key, and we should start thinking seriously about copyright registration. We do not have to reintroduce mandatory registration, but we do need to provide incentives to keep ownership information up to date and accessible, be it on a government-run registry or otherwise.²³⁰ Consequences for failing to

230. See Molly Shaffer Van Houwelling, *Author Autonomy and Atomism in Copyright Law*, 96 VA. L. REV. 549, 630–32 (2010) (discussing various mechanisms that can facilitate recordation and access to ownership information).

maintain copyright ownership information available will create demand for low-cost ownership data solutions.²³¹

Obviously, talking about registration immediately brings up the specter of violating the Berne Convention and its prohibition on formalities.²³² This Article will leave the discussion about remedy tweaking amounts to a prohibited “formality” for another day, but whether this is the case or not, Berne should not be an excuse for paralysis and inaction. The Berne Convention bans formalities principally because when it was negotiated, complying with the registration requirements that many countries imposed was overly onerous.²³³ But registration does not need to be onerous, and much can be done in ways that are consistent with Berne.²³⁴ We will never know what Berne exactly prescribes unless we test it. This Convention should not become the New Dogma, and Berne should not become the New Rome of the Copyright Church. If the Berne Convention does not allow us to do anything, then it should be reformed or abandoned.

The fact that we have created a system of ownership that does not require any record of ownership is one of the most serious copyright policy failures of our time. Copyright lasts longer than ever, it covers more works than ever, applies to more uses than ever, and the ways of using and reusing works are more complex and varied than ever. At the same time, there is no record of who owns what, at least not a universally accessible, usable, and reliable record. We want markets to function in this reality. They cannot. Developed markets cannot exist without a clear sense of who owns what, and markets for creative works are no exception. We would not expect real estate markets to function efficiently without land registration, and functioning securities markets without reliable records of securities’ ownership are hard to fathom. Why would we think that copyright is different? This ownership mess is not going to disappear spontaneously anytime soon, and it is not an American or Canadian or European problem.

231. Patry & Posner, *supra* note 32, at 1651.

232. See Berne Convention for the Protection of Literary and Artistic Works, Paris Act, July 24, 1971, 25 U.S.T. 1341, 828 U.N.T.S. 221, Art. 5(2) (“The enjoyment and the exercise of these rights shall not be subject to any formality”); see generally Stef Van Gompel, *Formalities in the Digital Era: An Obstacle or Opportunity?*, in GLOBAL COPYRIGHT: THREE HUNDRED YEARS SINCE THE STATUTE OF ANNE, FROM 1709 TO CYBERSPACE 395, 396 (L. Bently, U. Suthersanen & P. Torremans eds., 2010); Christopher Sprigman, *Reform(aliz)ing Copyright*, 57 STAN. L. REV. 485, 541 (2004).

233. See Van Gompel, *supra* note 232, at 418; Sprigman, *supra* note 232, at 544.

234. See Sprigman, *supra* note 232, at 554–68 (discussing Berne-compliant “new-style” formalities).

It is a global problem that affects every reading, watching, or listening person around the world.

As long as copyright law adheres to a strict permission-first rule and ignores owners' capacity to avoid the problem, the orphan works problem will appear to be a grand problem that requires grand solutions. But the political economy of copyright suggests that some grand solutions, such as radically reducing copyright terms, or reintroducing mandatory registration and renewal fees, are presently unrealistic, whereas more realistic grand solutions such as ECL may provide some immediate short-term fix, but might create persistent long-term side effects.

Copyright law was born out of revolt against the monopoly of the Stationers' Company over the book trade, and the role that it played, alongside the Crown and the Cross, in controlling the press and suppressing the free dissemination of knowledge, which the invention of print facilitated.²³⁵ We should keep this point in mind when we consider solutions to the orphan works problem, and we should be reluctant to adopt solutions that bring back a new version of the Stationers' Company, and entrust monopolies and bureaucracies as guardians of orphan works and gatekeepers of knowledge, as society moves from print to digital.

235. See sources cited *supra* note 8.

THE ORPHAN WORKS CHIMERA AND HOW TO DEFEAT IT: A VIEW FROM ACROSS THE ATLANTIC

Stef van Gompel[†]

TABLE OF CONTENTS

I.	INTRODUCTION.....	1347
II.	INTRODUCING THE CHIMERA: THE DIFFERENT MANIFESTATIONS OF THE PROBLEM	1351
	A. THE LION’S HEAD: MASS-DIGITIZATION PROJECTS.....	1351
	B. THE SERPENT’S HEAD: TRANSFORMATIVE AND DERIVATIVE USES.....	1356
	C. THE GOAT’S HEAD: SMALL-SCALE INCIDENTAL USES.....	1357
III.	A MULTIFACETED STRATEGY TO DEFEAT THE ORPHAN WORKS CHIMERA.....	1360
	A. TAMING THE LION: INTRODUCING COLLECTIVE LICENSING MECHANISMS FOR MASS-DIGITIZATION PROJECTS	1360
	B. KILLING THE SERPENT: APPROACHES BASED ON DILIGENT SEARCH REQUIREMENTS FOR TRANSFORMATIVE OR DERIVATIVE WORKS.....	1365
	C. RESTRAINING THE GOAT: PRACTICAL LEGAL SOLUTIONS FOR SMALL-SCALE INCIDENTAL USES	1373
IV.	CONCLUSION	1376

I. INTRODUCTION

For many years, scholars and policymakers around the world have attempted to answer the question of how best to address the problem of orphan works.¹ Orphan works are copyright protected works (or subject-

© 2012 Stef van Gompel.

† Stef van Gompel (LL.M. Amsterdam, LL.D. Amsterdam) is senior researcher at the Institute for Information Law (IViR) of the University of Amsterdam. You are invited to direct any comments, criticism, or ideas regarding this article to: vangompel@ivir.nl.

1. See generally P.B. HUGENHOLTZ ET AL., UNIV. OF AMSTERDAM INST. FOR INFO. LAW, THE RECASTING OF COPYRIGHT & RELATED RIGHTS FOR THE KNOWLEDGE ECONOMY 159–95 (2006) (reporting to the European Commission, DG Internal Market); Stef van Gompel, *Unlocking the potential of pre-existing content: How to address the issue of orphan works in Europe?*, 38 INT’L REV. INTELL. PROP. & COMPETITION L. 669 (2007); AGNIESZKA

matter protected by related rights),² the right owners of which cannot be identified or located by someone who wants to make use of such works in a manner requiring the right owners' consent. This may obstruct the entire process of clearing rights. Not being able to acquire the necessary permission from the right owners concerned makes it impossible to legally reproduce a work or communicate it to the public. Where the right owner cannot be found, even after a reasonably conducted search, a would-be user has only two choices: to use the work and bear the risk of an infringement claim, or to completely abandon his or her intention to use the work.³ The latter strategy forestalls a productive and beneficial use of the work. That is clearly not in the public interest, in particular where the right owner, if located, would not have objected to the use in question.

While different models have been suggested or adopted—varying from a limitation on remedy rule as proposed in the United States,⁴ to a system of mutual recognition of national solutions for orphan works as recently adopted in the European Union,⁵ to a specific exception or limitation as proposed in the United Kingdom⁶ and Germany,⁷ to a compulsory license to be obtained from a public authority as exists in Canada and other countries around the world,⁸ to an extended collective licensing scheme as applied in

VETULANI, THE PROBLEM OF ORPHAN WORKS IN THE EU: AN OVERVIEW OF LEGISLATIVE SOLUTIONS AND MAIN ACTIONS IN THIS FIELD 28–46 (2008) (report prepared for the European Commission, DG Information Society and Media, Unit E4: Digital Libraries and Public Sector Information).

2. Related (or neighboring) rights are rights related to or neighboring on copyright, such as the rights of performers, phonogram producers, film producers, and broadcasting organizations. The term “orphan works,” as used in this Article, will include subject matter protected by related rights (e.g., performances, phonograms, films and broadcasts).

3. U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 15 (2006), *available at* <http://www.copyright.gov/orphan/orphan-report-full.pdf>.

4. *Id.* at 12–13.

5. Directive 2012/28/EU, of the European Parliament and of the Council of 25 October 2012 on Certain Permitted Uses of Orphan Works, 2012 O.J. (L 299) 5 [hereinafter Orphan Works Directive].

6. *See* BRITISH SCREEN ADVISORY COUNCIL, COPYRIGHT AND ORPHAN WORKS: A PAPER PREPARED FOR THE GOWERS REVIEW 19 (2006), *available at* http://www.bsac.uk.com/files/copyright_orphan_works_paper_prepared_for_gowers_2006.pdf. The Gowers Committee recommended the proposal of the British Screen Advisory Council (BSAC) to the U.K. Government. *See* HM TREASURY, GOWERS REVIEW OF INTELLECTUAL PROPERTY 71 (2006), *available at* http://www.hm-treasury.gov.uk/d/pbr06_gowers_report_755.pdf.

7. AKTIONSBÜNDNIS URHEBERRECHT FÜR BILDUNG UND WISSENSCHAFT, BEDARF NACH EINER URHEBERRECHTSLÖSUNG FÜR VERWAISTE WERKE (2007), *available at* <http://www.urheberrechtsbuendnis.de/docs/verwaisteWerke.pdf>.

8. *See* Copyright Act, R.S.C. 1985, c. C-42, art. 77 (Can.). *See also* Chosakukenhou [Copyright Act], Act No. 48 of 1970, art. 67 (Japan); Jeojakgwonbeob [Copyright Act], Act

the Nordic countries⁹—they all fail to offer a complete answer to tackle the problem effectively. That is, rather than putting forward a general solution, each of the models provides relief only in specific circumstances.

To give an example, the models that are premised on the requirement to conduct a diligent search for the right owner—i.e., the proposed U.S. limitation on remedy rule, the E.U. system of mutual recognition, the proposed orphan works exception or limitation, and the Canadian statutory licensing scheme—are less suitable for mass-digitization projects or other projects where clearing rights on a work-by-work basis is not a viable option for cost-related reasons.¹⁰ The models that are based on collective rights management (e.g., the extended collective licensing scheme), on the other hand, are not really equipped for small-scale projects where one could easily perform a search for the right owner. Moreover, in practice, the latter models are useful only in those areas where collective rights management organizations actually operate.¹¹

This Article contends that the orphan works problem is a generic problem. It is not specific to a certain form of use, but has implications across the copyright spectrum. Yet, it manifests itself in different ways, depending on the type of use that is made of these works. The legal certainty that users of orphan works actually require is very much contingent on the

No. 9625, art. 50 (S. Kor.); Copyright Act, No 14 of 1957, INDIA CODE (1957), art. 31a. In the United Kingdom, where the Copyright Tribunal may give consent to the making of a copy of an “orphaned” recorded performance, a limited compulsory licensing scheme is also provided for. Copyright, Designs & Patents Act, 1988, c. 48, § 190 (U.K.).

9. JOHAN AXHAMN & LUCIE GUIBAULT, CROSS-BORDER EXTENDED COLLECTIVE LICENSING: A SOLUTION TO ONLINE DISSEMINATION OF EUROPE’S CULTURAL HERITAGE? 25 (2011), available at http://www.ivir.nl/publicaties/guibault/ECL_Europeana_final_report092011.pdf. In the United Kingdom, the Digital Economy Bill as originally introduced also included a provision enabling extended collective licensing for orphan works. Digital Economy Bill, 2009–10, H.L. Bill 1 cl. 42 (clause 42 was later renumbered as clause 43). Because it raised much controversy amongst right holders and photographers, in particular, see STOP43, <http://stop43.org> (last visited Sept. 10, 2012), the provision was eventually dropped. See Digital Economy Bill, 2009–10, H.L. Bill 55 (removing Clause 43 altogether). Currently, the Enterprise and Regulatory Reform Bill again contains a provision proposing new rules on the licensing of orphan works, including the possibility to obtain a license from a public authority and an extended collective licensing scheme. Enterprise and Regulatory Reform Bill, 2012–13, H.C. Bill 61 cl. 59.

10. Cf. MIREILLE VAN EECHOUDE ET AL., HARMONIZING EUROPEAN COPYRIGHT LAW: THE CHALLENGES OF BETTER LAWMAKING 278 (Info. Law Ser. 19, Kluwer Law International 2009).

11. See HUGENHOLTZ ET AL., *supra* note 1, at 184; Paul Goldstein, *The Quiet Revolution in American Copyright Law*, in PATENTS AND TECHNOLOGICAL PROGRESS IN A GLOBALIZED WORLD, MPI STUDIES ON INTELLECTUAL PROPERTY, COMPETITION AND TAX LAW 703, 706 (Wolrad Prinz zu Waldeck und Pyrmont et al. eds., 2009).

scale of use (large-scale versus incidental use) and the question of whether the works are used independently, are integrated, or are transformed into derivative works. As a consequence, the orphan works problem has different legal repercussions for different groups of users.

For this reason, this Article presents the orphan works problem as a Chimera, not in the sense of it being “an illusion or fabrication of the mind,”¹² but in the ancient mythological sense of a three-headed she-monster. As described in Homer and Hesiod, the Chimera was a fire-breathing monster with a lion’s head, a tail that ended in a serpent’s head, and a head of a goat in her middle.¹³ The metaphor of the Chimera enables the orphan works problem to be depicted as a generic problem with different manifestations, as represented by the three heads. Moreover, because chopping off only one of its heads will not immediately kill a multi-headed beast, the same metaphor illustrates that the only way to defeat the orphan works problem effectively is to pursue a multifaceted strategy. This forms the central thesis of this Article.

Accordingly, in examining how to address the orphan works problem, this Article argues that there is not one best approach. Rather, the problem requires different treatment depending on the type of use that is made of these works. Yet, there is an important contrast between Europe and the United States when it comes to adopting one model or the other. As will be seen, this is directly related to the difference between the two legal systems in terms of the financial damages that a user may incur when a right holder of an orphan work reappears. While this Article is predominantly written from a European viewpoint, it will draw a distinction between the legal systems in Europe and the United States to illustrate this point.

It must be emphasized that the orphan works problem does not occur where the consent of right owners is not required. This is the case, for example, where the act of reproduction or communication is covered by an exception or limitation. In Europe, an exhaustive list of exceptions and limitations can be found in Article 5 of the E.C. Directive on Copyright in the Information Society.¹⁴ This includes, *inter alia*, an exception in favor of archives, publicly accessible libraries, educational institutions, and museums

12. WEBSTER’S THIRD NEW INTERNATIONAL DICTIONARY OF THE ENGLISH LANGUAGE 389 (Philip Babcock Gove et al. eds., 1993).

13. HOMER, THE ILIAD 6.179–182 (Rodney Merrill trans., Univ. Mich. Press 2007); HESIOD’S THEOGONY 319–24 (Richard S. Caldwell trans., Focus Info. Grp. 1987).

14. Directive 2001/29/EC, of the European Parliament and of the Council of 22 May 2001 on the Harmonization of Certain Aspects of Copyright and Related Rights in the Information Society, 2001 O.J. (L 167) 10, 16–17.

to make specific acts of reproduction for non-commercial purposes.¹⁵ This allows European Union Member States to introduce a statutory exception to permit these institutions to make analogue or digital reproductions for purposes of preservation or restoration of works available in their collections. Similarly, in the United States, no permission from the copyright owner is required in cases where the fair use defense or any other statutory exception or limitation applies.¹⁶ Because the orphan works problem will not arise in cases covered by a specific exception or limitation, those situations will not be dealt with in this Article.¹⁷

This Article consists of three parts. Employing the metaphor of the Chimera, Part II systematically introduces the different ways in which the problem of orphan works manifests itself and describes the legal uncertainty that the different categories of users of orphan works experience. Next, Part III suggests addressing the orphan works problem by adopting a multifaceted approach that would provide adequate relief for the different categories of users of orphan works. Part IV concludes.

II. INTRODUCING THE CHIMERA: THE DIFFERENT MANIFESTATIONS OF THE PROBLEM

Although there are numerous ways in which the problem of orphan works is manifested, and generalizations are difficult to make, this Part identifies the three ways in which the orphan works problem materializes most prominently. For this purpose, it uses the Chimera to illustrate the different legal repercussions that the problem entails for different categories of users. Considering the main characteristics of the animals of which the Chimera is comprised, it depicts mass-digitization projects as the lion's head (Section II.A), transformative and derivative uses as the serpent's head (Section II.B), and small-scale incidental uses as the goat's head (Section II.C).

A. THE LION'S HEAD: MASS-DIGITIZATION PROJECTS

The topic that has drawn the most attention in the orphan works debate in recent years is mass-digitization. Mass-digitization projects can therefore

15. *Id.* Art. 5(2)(c).

16. *See* 17 U.S.C. §§ 107–122 (2010).

17. To the extent that a particular act of reproduction or communication can be brought under the fair use defense in the United States or an existing exception or limitation in Europe, this could bring sufficient relief to prospective users of orphan works. *See* Jennifer M. Urban, *How Fair Use Can Help Solve the Orphan Works Problem*, 27 BERKELEY TECH. L.J. 1379 (2012).

properly be characterized as the roaring lion of the Chimera. The lion is by far the loudest and noisiest of the three animals. Moreover, it is also one of the most dangerous and, except for the serpent, deadliest animals on earth.¹⁸ Perceived in the context of the present debate on orphan works, the latter characteristic quite well represents the legal uncertainty that the issue of orphan works creates for projects of mass-digitization.

In the European Union, almost the entire orphan works debate seems centralized around mass-digitization.¹⁹ Since 2005, orphan works have formed an integral part of the i2010: Digital Libraries initiative of the European Commission's DG Information Society.²⁰ This initiative aims to ensure that a full catalog of cultural and scientific materials available in libraries, museums, and archives across Europe (including literature, art, films, and photos) is digitized and made available online. In the 2006 Commission Recommendation on the digitization and online accessibility of cultural material and digital preservation, the Commission called upon the Member States to create mechanisms that would facilitate the use of orphan works and to promote the availability of lists of known orphan works and works in the public domain;²¹ but few Member States responded to this call.²² Therefore, the Commission took the matter in its own hands and put forward a proposal for a Directive on certain permitted uses of orphan works, which has recently been adopted.²³ This Directive is aimed entirely at enabling the digitization of and ensuring "the lawful, cross-border online

18. See David Pegg, *The 25 Most Dangerous Animals in the World*, LIST25 (Nov. 23, 2011), <http://list25.com/the-25-most-dangerous-animals-in-the-world>. The African lion is number fourteen on the list of twenty-five most dangerous animals in the world, as compiled by List25. By contrast, serpents such as the boomslang and the carpet viper are at numbers thirteen and seven, respectively. Unsurprisingly, goats do not appear on this list.

19. See, e.g., VAN EECHOUDE ET AL., *supra* note 10, at 263–65.

20. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 'i2010: Digital Libraries'*, at 6–7, COM (2005) 465 final (Sept. 30, 2005). See also *Digital Agenda for Europe: Digital Libraries Initiative*, EUROPE'S INFORMATION SOCIETY THEMATIC PORTAL, http://ec.europa.eu/information_society/activities/digital_libraries/index_en.htm (last visited Aug. 29, 2012).

21. Commission Recommendation of 24 August 2006 on the Digitization and Online Accessibility of Cultural Material and Digital Preservation 2006/585/EC, arts. 6(a)(c), 2006 O.J. (L 236) 28, 30.

22. See *Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions, Europe's Cultural Heritage at the Click of a Mouse: Progress on the Digitisation and Online Accessibility of Cultural Material and Digital Preservation Across the EU*, at 6–7, COM (2008) 513 final (Aug. 11, 2008). See also *Commission Staff Working Paper accompanying the Commission Recommendation on the Digitisation and Online Accessibility of Cultural Material and Digital Preservation*, at 26–27, SEC (2011) 1274 final (Oct. 27, 2011).

23. Orphan Works Directive, *supra* note 5.

access to orphan works contained in ... particular cultural heritage institutions.”²⁴

Likewise, the U.S. Copyright Office adopted a preliminary analysis and discussion document on legal issues in mass digitization in October 2011.²⁵ In contrast to the U.S. Copyright Office’s report on orphan works of 2006, which treats the problem in a generic way, this document discusses the issue of orphan works entirely in the context of mass-digitization.²⁶ The present discussion paper invites interested parties to start a debate on possible approaches—voluntary or legislative—to tackling the orphan works problem, specifically for projects of mass-digitization.²⁷

One reason why the U.S. Copyright Office adopted this preliminary analysis and discussion document is, obviously, the rejection by the U.S. District Court for the Southern District of New York of the class action lawsuit brought by the Authors Guild and others against Google for the mass digitization of books in several large U.S. libraries.²⁸ The Google Books Settlement raised controversy, inter alia, because it would grant Google a monopoly over out-of-print works and orphan works covered by the class action.²⁹ In rejecting the proposed settlement, the court ruled that the “question of who should be entrusted with guardianship over orphan books, under what terms, and with what safeguards are matters more appropriately decided by Congress than through an agreement among private, self-interested parties.”³⁰

Orphan works indeed present a major obstacle to projects of mass-digitization. That is due in particular to the scale in which pre-existing works are digitized and made available. Although the size of the problem is difficult

24. See Stef van Gompel, *The Commission’s Proposal for a Directive on Orphan Works*, KLUWER COPYRIGHT BLOG (June 14, 2011), <http://kluwercopyrightblog.com/2011/06/14/the-commission’s-proposal-for-a-directive-on-orphan-works>; Stef van Gompel, *Het richtlijnvoorstel verweesde werken: Een kritische beschouwing*, 35 AMI TIJDSCHRIFT VOOR AUTEURS-, MEDIA- & INFORMATIERECHT 205, 207–08 (2011).

25. U.S. COPYRIGHT OFFICE, LEGAL ISSUES IN MASS DIGITIZATION: A PRELIMINARY ANALYSIS AND DISCUSSION DOCUMENT (2011), available at http://www.copyright.gov/docs/massdigitization/USCOMassDigitization_October2011.pdf.

26. *Id.* at 25–29.

27. *Id.* at i, 1–4.

28. Authors Guild v. Google Inc., 770 F. Supp. 2d 666 (S.D.N.Y. 2011).

29. See, e.g., James Grimmelman, *How to Fix the Google Book Search Settlement*, 12 J. INTERNET L. 1, 14 (2009); Randal C. Picker, *The Google Book Search Settlement: A New Orphan-Works Monopoly?*, 5 J. COMPETITION L. & ECON. 383 (2009); Pamela Samuelson, *The Google Book Settlement as Copyright Reform*, 2011 WIS. L. REV. 479, 511 (2011).

30. *Authors Guild*, 770 F. Supp. 2d at 677.

to quantify,³¹ and precise figures are impossible to give,³² more and more studies show that the number of orphan works available in the collections of libraries, archives, and museums is significant.³³ Estimates at the high end of the range are provided for, inter alia, by the British Library, which suggests that at least 40% of the works in their collections are orphaned.³⁴ More conservative estimates indicate that, in Europe, about three million books (i.e., 13% of the total number of in-copyright books) and between 129,000 and 225,000 cinematographic works are orphan works.³⁵

Thus, if a party engages in the large-scale digitization and communication to the public of the works held by libraries, museums, or archives, and, due to the orphan status of many works, is unable to secure prior authorization from the right owners concerned, it risks exposing itself to claims for financial damages. For various mass-digitization projects these risks may well become prohibitively high, as “many small claims can add up to a considerable total.”³⁶

This is certainly the case in the United States, where users face the risk of becoming liable for payment of statutory damages of up to \$150,000 for each willfully committed infringement.³⁷ This can happen if the user had constructive knowledge of the infringing nature of his or her conduct.³⁸ This

31. See U.S. COPYRIGHT OFFICE (2006), *supra* note 3, at 92; HUGENHOLTZ ET AL., *supra* note 1, at 167.

32. An important reason is that the size of the orphan works problem is inversely proportional to the search that is conducted to identify and locate the right owners of the works. See Van Gompel, *Het richtlijnvoorstel verweesde werken: Een kritische beschouwing*, *supra* note 24, at 205.

33. See, e.g., DENISE TROLL COVEY, ACQUIRING COPYRIGHT PERMISSION TO DIGITIZE AND PROVIDE OPEN ACCESS TO BOOKS 16–21, 24–35 (2005), available at <http://www.clir.org/pubs/reports/pub134/pub134col.pdf> (establishing “that identifying and locating current copyright owners, particularly of older books, is a difficult, time-consuming, hit-or-miss, sometimes futile process”); NAOMI KORN, JOINT INFO. SYS. COMM., IN FROM THE COLD: AN ASSESSMENT OF THE SCOPE OF “ORPHAN WORKS” AND ITS IMPACT ON THE DELIVERY OF SERVICES TO THE PUBLIC 5 (2009), available at http://sca.jiscinvolve.org/wp/files/2009/06/sca_colltrust_orphan_works_v1-final.pdf; ANNA VUOPALA, EUR. COMM’N, DG INFO. SOC’Y & MEDIA: UNIT E4 ACCESS TO INFO., ASSESSMENT OF THE ORPHAN WORKS ISSUE AND COSTS FOR RIGHTS CLEARANCE 5 (2010), available at http://ec.europa.eu/information_society/activities/digital_libraries/doc/reports_orphan/anna_report.pdf.

34. See BRITISH LIBRARY, COPYRIGHT FOR EDUCATION AND RESEARCH: GOLDEN OPPORTUNITY OR DIGITAL BLACK HOLE? 1 (2009), available at <http://pressandpolicy.bl.uk/imagelibrary/downloadMedia.ashx?MediaDetailsID=564>.

35. VUOPALA, *supra* note 33, at 5, 43.

36. VAN EECCHOUD ET AL., *supra* note 10, at 271.

37. 17 U.S.C. § 504(c) (2010).

38. See, e.g., *Island Software & Computer Serv., Inc. v. Microsoft Corp.*, 413 F.3d 257, 264 (2d Cir. 2005).

is particularly risky for users of orphan works, who are likely to possess such knowledge considering that they use works knowing that their right owners are unknown or untraceable.³⁹ However, the court shall remit statutory damages if it finds that the infringer is a nonprofit educational institution, library, or archive that believed and had reasonable grounds for believing that its copying was a fair use.⁴⁰ Still, risk-averse institutions may well avoid relying on the fair use defense, as it does “not offer absolute protection from allegations of infringement.”⁴¹ The chance of being exposed to statutory damages simply poses too great a risk. A perfect illustration is the example given by Samuelson and Wheatland, who calculate that, in the *Author’s Guild v. Google, Inc.* case, even “an award of the statutory minimum of \$750 per book would yield approximately \$4.5 billion in liability” for Google.⁴²

By contrast, in Europe, damages are ordinarily based on the actual losses incurred by the infringement.⁴³ Unlike in the United States, the damages recoverable are compensatory and not punitive in nature. Therefore, if a user eventually is held liable for copyright infringement, he or she generally is obligated to pay no more than the actual damages suffered by the right owner, which is often comparable to the amount the user would have had to pay in case a license had been obtained. Accordingly, in Europe, users of orphan works do not have the Damocles sword of large monetary awards hanging over them.⁴⁴ This allows users to reserve a certain amount of money to pay a right owner should he or she eventually come forward.

39. Cf. Memorandum of Amicus Curiae The Internet Archive in Opposition to Settlement Agreement at 3–4, *Authors Guild v. Google Inc.*, No. 05-CV-08136-DC (S.D.N.Y. Sept. 8, 2009).

40. See 17 U.S.C. § 504(c)(2) (2010).

41. U.S. COPYRIGHT OFFICE (2011), *supra* note 25, at 19.

42. Pamela Samuelson & Tara Wheatland, *Statutory Damages in Copyright Law: A Remedy in Need of Reform*, 51 WM. & MARY L. REV. 439, 491 (2009).

43. See Directive 2004/48/EC, of the European Parliament and of the Council of 29 April 2004 on the Enforcement of Intellectual Property Rights (Corrigendum), art. 13(1), 2004 O.J. (L 195) 16, 23 (instructing Member States to ensure the availability of “damages appropriate to the actual prejudice suffered by [the right holder] as a result of the infringement”). While this is also the general rule under the common law of the U.K., see L. BENTLY & B. SHERMAN, *INTELLECTUAL PROPERTY LAW* 1117 (3d ed. 2009), additional damages are available under the Copyright, Designs & Patents Act, 1988, c. 48, § 97(2) (U.K.). Their award, however, is the exception rather than the rule. If they nevertheless are awarded, the level of additional damages typically has been modest. See Christina Michalos, *Copyright and Punishment: The Nature of Additional Damages*, 22 EUR. INTEL. PROP. REV. 470, 473 (2000).

44. To have the sword of Damocles hanging over you means that, despite apparent well-being, you are facing an ever-present threat. The expression alludes to the legend of Damocles, a courtier to King Dionysius I of Syracuse. According to the legend, Damocles so frequently praised the power and apparent happiness of his king that the latter invited him to

For most mass-digitization projects, however, it is difficult to estimate expected future claims resulting from the use of orphan works. In practice, it is not always clear how large monetary reserves need to be. The legal uncertainty that this entails may consequently be too high to risk using these works.⁴⁵ Moreover, for libraries, museums, and archives, the prospect of being held liable for copyright infringement is very unnerving. Since these institutions are often nonprofit and funded by public money, they cannot tolerate getting a reputation of being large-scale infringers of copyright-protected works.⁴⁶

On the other hand, the legal uncertainty that the orphan works problem creates for mass-digitization projects does not extend beyond financial loss. If right owners that reappear seek injunctive relief by prohibiting further use of their works, the projects usually do not run high risks, as the works can simply be removed from their online databases.⁴⁷

B. THE SERPENT'S HEAD: TRANSFORMATIVE AND DERIVATIVE USES

The problem of orphan works manifests itself somewhat differently in cases of transformative and derivative uses. Examples are motion pictures and documentary films that incorporate pre-existing material, such as photographs and film footage. When a filmmaker comes across a work that he or she wants to integrate into a movie, perhaps because it has unique historical or cultural merit or simply because it fits the theme or storyline, the filmmaker may be legally liable if the right owner of the work cannot be identified or located to ask for permission.

In contrast to mass-digitization projects, transformative or derivative uses of orphan works appear to carry less financial risk. Transformative and derivative uses usually concern reutilization on a small scale. Except for a situation in which a derivative work incorporates a myriad of orphan works, it seems feasible to estimate the financial liability that may arise when using a work without being able to obtain the prior authorization of its right owner.⁴⁸ As observed in Section II.A, *supra*, however, the financial risks for creative re-users in the United States are likely to be higher than for those in Europe,

a banquet and seated him under a sword hung by a single horse-hair, so as to point out to him the fears and worries that kingship brings with it.

45. See KORN, *supra* note 33, at 28–29.

46. *Id.* at 21 (quoting Clifford Lynch, Director of the Coalition for Networked Information).

47. VAN EECHOUDE ET AL., *supra* note 10, at 272.

48. See KORN, *supra* note 33, at 6, 20 (indicating as one of the key research findings that “the most common method for managing Orphan Works is the adoption of a risk managed approach (average of 60%).”).

given the availability of statutory damages with a punitive character under U.S. copyright law.

Having said that, injunctions are by far the most serious cause of legal uncertainty for persons making transformative or derivative uses of orphan works. Where an orphan work has been incorporated into a derivative work, a resurfaced right owner may file suit against the user and claim—in addition to awards for damages—an injunction preventing the continued use of his or her work.⁴⁹ This may have grave consequences. For example:

If an orphan work is used in a documentary film . . . and the further use of this work would be forbidden by a reappeared right owner seeking injunctive relief, the documentary filmmaker may run the risk that the orphan work needs to be removed (which may render the entire film worthless) or, in the worst case scenario, that the entire film needs to be taken from the market.⁵⁰

For this reason, transformative and derivative uses of orphan works can best be characterized as the serpent's tail of the Chimera. Serpents are silent, yet very effective killers. Slowly but successfully, their one bite can kill an entire living being. This quite well resembles the legal uncertainty caused by the availability of injunctive relief. If a transformative or derivative user of an orphan work is confronted with an injunction filed by a reappeared right owner, he or she may at worst witness the entire project being brought to an end. Hence, there is much more at stake in case an orphan work is integrated or transformed into a derivative work than when it has simply been reproduced and made online available to the public without any transformation of its content, as in the majority of mass-digitization projects.⁵¹

C. THE GOAT'S HEAD: SMALL-SCALE INCIDENTAL USES

A third group of uses involves certain small-scale incidental uses. Given the relatively little harm that the orphan works problem causes for this category of uses, at least in Europe, it will be represented by the goat, which is by far the most innocent of the three animals comprising the Chimera.

One example of small-scale incidental use is the use of orphaned film footage in television-programs. Broadcasters generally clear the rights for the

49. See U.S. COPYRIGHT OFFICE (2006), *supra* note 3, at 13.

50. VAN EECHOUDE ET AL., *supra* note 10, at 272.

51. For this reason, the U.S. Copyright Office has suggested introducing a limitation on injunctive relief for situations where an orphan work is incorporated into a derivative work, but not for cases in which orphan works are used without transformation of the content. See U.S. COPYRIGHT OFFICE (2006), *supra* note 3, at 120.

majority of works included in a broadcast, e.g., via agreements with collective rights management organizations, film studios, and other right owners. Sometimes, however, they may want to include works that are not covered by these agreements, such as archived materials, the right owners of which are not always known. In such cases, while it is impossible to appropriately clear all the rights, broadcasters may nonetheless decide to use the work, as the risks involved are not particularly high. First, if licenses have been properly arranged for the majority of works, the chances that an unrepresented right owner reappears seem limited. Moreover, by the time that the right owner comes forward, the television-program will most probably have already been broadcast. This would render filing an injunction claim useless, given that it would only prohibit future occasions. Consequently, should a right owner come forward, he or she can claim monetary awards at best. Broadcasters can reserve monetary funds to recompense resurfaced right owners for this purpose. In Europe, this already occurs in practice.⁵²

Another example is the inclusion of orphaned images in CD or DVD booklets. Entertainment companies will most likely have cleared the rights in the layout, the artwork, and the design of such booklets and certainly in the recorded music and film works. Nevertheless, in practice it may happen that, incidentally, for some works the necessary permission cannot be arranged for the simple reason that the copyright holders are unknown or untraceable. For example, in the case of photos of “old” artists of whom only a handful of photos are available, identifying who owns the rights in the photographic

52. Broadcasters that incidentally use an orphan work in a television broadcast often indicate this by including a legal disclaimer in the closing credits inviting a reappearing right holder to contact them to get reimbursement for the use made. *See infra* Section III.C. This suggests that they have reserved some funds for this purpose. The Dutch broadcasting organization NPS for example sometimes uses the following disclaimer:

De NPS heeft haar uiterste best gedaan om alle rechthebbenden van door ons gebruikt archiefmateriaal te achterhalen. Indien een persoon of bedrijf echter kan aantonen dat haar toestemming m.b.t. door ons gebruikt archiefmateriaal was vereist, dan verzoeken wij deze om binnen twee maanden na uitzending contact met ons op te nemen.

Translated from the original Dutch:

The NPS has done her utmost best to trace all right holders of the archival materials used. If a person or company can nevertheless demonstrate that her permission was required for our usage of the archival materials, then we kindly request this person or company to contact us within two months after the broadcast.

works may not always be an easy task.⁵³ Here again, the legal uncertainty of using such photos without prior authorization may be fairly limited. A reappeared right owner may of course seek injunctive relief. However, if a risk assessment shows that the chances of this are limited, a record company may well bear such risk, because changing a booklet of a CD or DVD does not incur prohibitive costs. Furthermore, if a right owner comes forward, he or she may well choose to settle for monetary awards, rather than start a lengthy and costly legal proceeding to put an end to this use.⁵⁴ It could therefore prove satisfactory to create an escrow account for compensating reappeared right holders.⁵⁵

Admittedly, in both examples, the use in question infringes the exclusive rights of copyright owners. Clearly this must not be understood as *carte blanche* for users to use a work without undertaking a diligent search to find the right owner in order to ask for consent and arrange payment for the use made.⁵⁶ Although identifying and locating right holders may well be a difficult and time-consuming process,⁵⁷ this does not exempt prospective users from spending sufficient time and resources in seeking permission.⁵⁸ The fact that a right owner of a work is unknown or unlocatable should not, at the same time, be an absolute impediment to the reutilization of such work. There ought to be a balance between the two. The two cases show that, within the confines of current legislation and without sacrificing legal certainty for users or prejudicing the legitimate interests of right holders, users in Europe may already find ways to incidentally use orphan works.

In the United States, on the other hand, it would be impossible for users to make small-scale incidental uses of orphan works without facing the risk

53. See, e.g., VAN EECHOUDE ET AL., *supra* note 10, at 263 (stating that “it is more likely that works become orphaned as they grow older”).

54. See WILLIAM M. LANDES, *An Empirical Analysis of Intellectual Property Litigation: Some Preliminary Results*, 41 HOUS. L. REV. 749, 753–57, 767 (2004) (suggesting that, from an economic point of view, “an overall increase in the costs of going to trial relative to the costs of settling should reduce the number of cases filed, terminated, and tried.”).

55. See COMITÉ DES SAGES, THE NEW RENAISSANCE, 18 ¶ 5.3.2.2(vi) (2011), available at http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/final_report_cds.pdf (recommending an escrow account).

56. In 2006, certain groups of right holders, including photographers, illustrators, and graphic artists, expressed the fear that their works would unjustly be labeled as orphan works and used without prior authorization under the proposed limitation on remedy rule in the United States. See *Orphan Works: Proposals for a Legis. Solution: Hearing before the Subcomm. on Intell. Prop. of the Comm. on the Judiciary of the U.S. Senate*, 109th Cong. 7–8, 10–12, 85–87, 95–99 (2006) (statements of Victor Perlman and Brad Holland).

57. See COVEY, *supra* note 33, at 21, 24, 35.

58. See HUGENHOLTZ ET AL., *supra* note 1, at 167.

of becoming liable for payment of statutory damages.⁵⁹ This may constitute a real obstacle for users hoping to engage with secondary uses of orphan works similar to the ones discussed here. As will be argued in Section III.C, *infra*, addressing this matter would require the United States, but not Europe, to take action.

III. A MULTIFACETED STRATEGY TO DEFEAT THE ORPHAN WORKS CHIMERA

The analysis in Part II, *supra*, reveals that the orphan works problem has different legal implications for different types of uses. Tackling it effectively would therefore require taking action on multiple fronts. Differentiating between the three manifestations of the problem, as metaphorically represented by the animals that comprise the Chimera, this Part describes the multifaceted strategy that could be used to address the orphan works problem. The strategy is not aimed at slaying the Chimera entirely, but rather at taming the lion, killing the serpent, and restraining the goat. It suggests tackling the problem by introducing collective licensing mechanisms for mass-digitization projects (Section III.A), adopting approaches based on diligent search requirements for transformative and derivative uses (Section III.B), and relying on practical legal solutions for small-scale incidental uses (Section III.C).

A. TAMING THE LION: INTRODUCING COLLECTIVE LICENSING MECHANISMS FOR MASS-DIGITIZATION PROJECTS

As observed in Section II.A, *supra*, the legal uncertainty that the problem of orphan works creates for mass-digitization projects is directly related to the large amount of works that are being reutilized. Due to the scale of use, clearing rights on an individual basis would be too costly for most of these projects. Legislative models based on a requirement to conduct a diligent search for the right owners of a work therefore would not offer a sound solution for such projects.⁶⁰ A better approach would be to adopt a model

59. See *infra* Section III.C.

60. For this reason, the Orphan Works Directive, see *supra* note 5, is highly criticized by European scholars. It aims at facilitating the digitization and online access of materials stored in libraries, museums, and archives through a system of mutual recognition of national solutions for orphan works that are premised on the requirement to conduct a diligent search for the right owner. See Van Gompel, *Het richtlijnvoorstel verveesde werken: Een kritische beschouwing*, *supra* note 24, at 216–17; Allard Ringnalda, *De voorgestelde richtlijn Verveesde Werken: op naar een Europese internetbibliotheek?*, 27 *INTELLECTUELE EIGENDOM EN RECLAMERECHT* 387, 391 (2011); See generally Robin Kerremans, *A critical view on the European draft directive for orphan works*, 2 *QUEEN MARY J. INTELL. PROP.* 38 (2012).

based on collective licensing.⁶¹ Collective licensing has the advantage of enabling users to obtain a license to use a multitude of works with a single transaction. Accordingly, right holders need not be traced and asked for permission on an individual basis. Collective agreements thus provide an efficient tool for mass-reutilization projects, which by definition require rights clearance for a myriad of works.⁶²

The collective licensing of rights also alleviates the problem of orphan works. If a collective rights management organization (“CRMO”) has been established and it represents a significant part of right holders in a given field, there is a reasonable likelihood that the particular right owner the user is looking for will also be represented.⁶³ In that case, users face fewer difficulties in finding the right owners whose works they intend to use. Yet, if a right owner is not represented by that CRMO, a user may still face considerable uncertainties.⁶⁴

There are different legal techniques, however, to ensure that CRMOs can issue fully covering licenses to users of copyrighted works.⁶⁵ One such technique is extended collective licensing (“ECL”), which is applied in various sectors in Denmark, Finland, Norway, Sweden, and Iceland.⁶⁶ The most characteristic feature of the ECL scheme is that the law gives extended effect to a freely negotiated collective licensing agreement concluded

61. VAN EECHOUDE ET AL., *supra* note 10, at 278.

62. Allard Ringnald, *Orphan Works, Mass Rights Clearance, and Online Libraries: The Flaws of the Draft Orphan Works Directive and Extended Collective Licensing as a Solution*, 8 *MEDIEN UND RECHT INT'L* 3 (2011).

63. Stef van Gompel & P.B. Hugenholtz, *The Orphan Works Problem: The Copyright Conundrum of Digitizing Large-Scale Audiovisual Archives, and How to Solve it*, 8 *POPULAR COMM.* 61, 65 (2010).

64. See Jean-François Debarnot, *Les droits des auteurs des programmes du fonds de l'INA exploités sur son site Internet*, 232 *LÉGIPRESSE* 93, 93–94 (2006) (indicating that the collective agreement concluded in France between the National Audiovisual Institute (INA) and five French CRMOs, authorizing the use of the entire CRMOs' audiovisual and sound catalog available in INA's archive for any mode of exploitation, does not cover the repertoire of right holders who are not members of any of the five CRMOs and that, as a result, INA still needs to identify and locate these, perhaps unknown, right holders to clear the rights of the works not covered by the agreement).

65. See generally Van Gompel (2007), *supra* note 1, at 686–91; VAN EECHOUDE ET AL., *supra* note 10, at 278–82 (discussing four legal techniques enabling CRMOs to issue “blanket”—i.e., fully covering—licenses: extended collective licensing, legal presumption of representation, contracts with indemnity clauses, and mandatory collective rights management).

66. See Tarja Koskinen-Olsson, *Collective management in the Nordic countries*, in *COLLECTIVE MANAGEMENT OF COPYRIGHT AND RELATED RIGHTS* 283, 283 (Daniel Gervais ed., Kluwer Law International 2nd ed. 2010); AXHAMN & GUIBAULT, *supra* note 9, at 25–60.

between a representative CRMO and a user or a group of users.⁶⁷ While the collective licensing system is premised on a voluntary transfer of rights from right holders to a CRMO, the extended effect ensures that the collective licensing agreement to which the ECL rules apply also covers the repertoire of right holders who are not members of the CRMO.⁶⁸ A condition is that the CRMO is sufficiently representative, meaning that a substantial number of right holders in a given category must have entrusted it with the exploitation of their rights.⁶⁹

Since the ECL provisions extend the operation of particular collective licensing agreements to all right holders in the given field, irrespective of whether they are domestic or foreign, the agreements also automatically cover the heirs of deceased right holders and all right holders that are unknown or untraceable.⁷⁰ This greatly facilitates the clearance of rights, since a user may obtain a license to use all works covered by the license without the risk of infringing the rights of right owners who otherwise would not be represented. In fact, the rationale of the system of ECL has always been to facilitate the licensing in case of massive uses for which it would be impossible for users to clear all the necessary rights.⁷¹

To protect the interests of right holders who have not mandated the CRMO to exploit the rights on their behalf, the law in the Nordic countries often contains two possibilities for right holders to stop (fully) participating in the ECL scheme: they have the right to claim individual remuneration, or to “opt out” from the ECL altogether.⁷² This allows non-represented right holders to negotiate a higher licensing fee and/or to enforce their copyrights against an exploiter on an individual basis.⁷³

67. Daniel Gervais, *Collective Management of Copyright: Theory and Practice in the Digital Age*, in COLLECTIVE MANAGEMENT OF COPYRIGHT AND RELATED RIGHTS 1, 21 (Daniel Gervais ed., Kluwer Law International 2nd ed. 2010).

68. See AXHAMN & GUIBAULT, *supra* note 9, at vii (referring to these non-represented right holders as “outsiders”).

69. See Henry Olsson, *The Extended Collective License as Applied in the Nordic Countries*, KOPINOR 25TH ANNIVERSARY INT’L SYMPOSIUM, § 6.2 (Oct. 3, 2010), available at <http://www.kopinor.no/en/copyright/extended-collective-license/documents/the-extended-collective-license-as-applied-in-the-nordic-countries>.

70. Gervais, *supra* note 67, at 21. See also Marco Ricolfi, *Digital Libraries in the Current Legal and Educational Environment: a European Perspective*, in GLOBAL COPYRIGHT: THREE HUNDRED YEARS SINCE THE STATUTE OF ANNE, FROM 1709 TO CYBERSPACE 216, 223–24 (Lionel Bently et al. eds. 2010).

71. Olsson, *supra* note 69, § 3.

72. Van Gompel & Hugenholtz, *supra* note 63, at 66. For a description of these exit options under the ECL schemes of the Nordic countries, see Olsson, *supra* note 69, § 6.4.

73. It has been argued that the process of “opting out” should be fairly simple and straightforward to prevent it from becoming a de facto formality prohibited by art. 5(2)

From the perspective of users, the system of ECL may offer a valuable solution to the orphan works problem. The extended effect of the ECL assures that a collective licensing agreement is fully covering and “applies to all right holders in the given field (except to those who have explicitly opted out from the system).”⁷⁴ Consequently, under the system of ECL, users no longer have to anticipate unexpected claims from non-represented right holders. Therefore, an ECL agreement “provides re-users of existing works with a considerable extent of legal certainty that they require.”⁷⁵

From the perspective of right holders, on the other hand, introducing an ECL would be quite a radical solution. Therefore, if a system like this would be established,

[it] should only be applied in cases where there is a clear public interest at stake. Examples include the exploitation of past archive productions of public broadcasting organisations for on demand services; or the exploitation of copyrighted works included in the collection of archives, museums, libraries or educational institutions for specific purposes such as public exhibition, private studying, teaching or scientific research.⁷⁶

An often-claimed disadvantage of the ECL model is that it requires an upfront payment of royalties that cannot be properly distributed to the rightful copyright owners as long as they remain unknown or untraceable.⁷⁷ The model therefore is said to come at a high price.⁷⁸ For mass-digitization projects, however, it seems that the additional costs of a license that covers

Berne Convention. See Mihály Ficsor, *Collective Management of Copyright and Related Rights from the Viewpoint of International Norms and the Acquis Communautaire*, in COLLECTIVE MANAGEMENT OF COPYRIGHT AND RELATED RIGHTS 29, 62 (Daniel Gervais ed., Kluwer Law International 2nd ed. 2010). But see Gervais, *supra* note 67, at 22–27; AXHAMN & GUIBAULT, *supra* note 9, at 46; Stef VAN GOMPEL, FORMALITIES IN COPYRIGHT LAW: AN ANALYSIS OF THEIR HISTORY, RATIONALES AND POSSIBLE FUTURE 209–11 (Info. Law Ser. 23, Kluwer Law International 2011) (contending that the “opt out” mechanism under the ECL system does not constitute a formality prohibited by the Berne Convention).

74. VAN EECHOUDE ET AL., *supra* note 10, at 280.

75. *Id.* at 280. But see Van Gompel & Hugenholtz, *supra* note 63, at 66 (indicating that the Nordic system of ECL, “albeit highly practical and attractive, cannot provide complete certainty to prospective users” because right holders who opt out are no longer covered by it). To enhance certainty, however, the names and contact details of right holders opting out of the ECL could be recorded in a database. See VAN EECHOUDE ET AL., *supra* note 10, at 280 n.859.

76. HUGENHOLTZ ET AL., *supra* note 1, at 194.

77. See *Collective Management of Copyright: Solution or Sacrifice? Panel: Collective Licensing for Digitizing Analog Materials*, 34 COLUM. J.L. & ARTS 747, 747–48 (2011).

78. Katharina de la Durantaye, *How to Build an Orphanage, and Why*, 2 J. INTELL. PROP. INFO. TECH. & E-COMMERCE L. 226, 228–29 (2011).

orphan works do not outweigh the benefits that collective licensing has over individual rights clearance. A study by the British Library shows that searching for right holders on a work-by-work basis is very costly in terms of staff time.⁷⁹ Furthermore, the non-distributable royalties accumulating in the ECL system could perhaps be used for other purposes, e.g., “for cultural aims [including the funding of new mass-digitization projects], or for the creation of databases of rights management information that will prevent future works from becoming ‘orphaned.’”⁸⁰

Difficulties may arise, moreover, in the practical implementation of an ECL regime. Since the success of it fully depends on the conclusion of contracts between users and CRMOs that represent a sufficient number of right holders, CRMOs should already be operating in those fields where the orphan works problem is most pressing. This is currently not the case. In Europe, collective rights management is underdeveloped in the photographic and audiovisual fields, in particular. Right owners in those areas often are cautious to participate in collective licensing schemes and prefer to manage their rights individually. This may prevent the ECL model from becoming a successful solution to the problem of orphan works.⁸¹

Collective exercise of rights is conceivably much more poorly developed in the United States than in Europe.⁸² The recently rejected Google Books Settlement can illustrate this. In the absence of relevant CRMOs, the settlement would have required the creation of a Books Rights Registry. This proposed Registry would collect the royalties that Google had to pay for the digitization, make the books available to the public, and distribute royalties among relevant right owners. Thus, the Registry effectively would be a new CRMO for the administration of the rights of authors whose works were

79. BARBARA STRATTON, SEEKING NEW LANDSCAPES: A RIGHTS CLEARANCE STUDY IN THE CONTEXT OF MASS DIGITISATION OF 140 BOOKS PUBLISHED BETWEEN 1870 AND 2010, 5 (2011), available at <http://www.arrow-net.eu/sites/default/files/Seeking%20New%20Landscapes.pdf> (indicating as one of the key findings that “[o]n average it took 4 hours per book to undertake a ‘diligent search.’”).

80. HUGENHOLTZ ET AL., *supra* note 1, at 194–95. See also Van Gompel, *supra* note 1, at 701. To prevent the non-distributable royalties in an ECL scheme from accumulating too heavily, CRMOs could be ordered, in fixing the price for a license, to take into account the number of unrepresented right holders in the particular field and/or the estimated number of orphan works that would likely be used by the user to whom the license is granted.

81. VAN EECHOUDE ET AL., *supra* note 10, at 279.

82. See Goldstein, *supra* note 11, at 706 (indicating that, because of “the narrowness of their focus,” copyright collecting societies in the United States offer “a comparably less satisfactory solution” than in other countries around the world). For an overview of collective licensing in the United States, see Glynn Lunney, *Copyright Collectives and Collecting Societies: The United States Experience*, in COLLECTIVE MANAGEMENT OF COPYRIGHT AND RELATED RIGHTS 29, 62 (Daniel Gervais ed., Kluwer Law International 2nd ed. 2010).

covered by the settlement.⁸³ Interestingly, because the settlement was negotiated between the parties in a class action lawsuit, it would have bound all copyright owners in the relevant class, had the court approved it. This would have given the Google Books Settlement an extended effect similar to the ECL system.⁸⁴

B. KILLING THE SERPENT: APPROACHES BASED ON DILIGENT SEARCH REQUIREMENTS FOR TRANSFORMATIVE OR DERIVATIVE WORKS

A collective licensing system as discussed in Section III.A, *supra*, would not offer a suitable model for enabling transformative or derivative uses of orphan works. In comparison with mass-digitization projects, derivative works usually incorporate pre-existing works at a relatively small scale. It should cause no difficulties for transformative users, therefore, to perform a good faith search for the right owners of these works. Accordingly, an approach based on a diligent search requirement could provide ample relief for this category of users. Because injunction claims are by far the most serious cause of legal uncertainty for transformative users,⁸⁵ it seems elementary that a legislative model provides bona fide users with adequate legal safeguards to properly eliminate this uncertainty.

In practice, there are various orphan-works-tailored legislative models that are premised on a diligent search.⁸⁶ One example is the limitation on remedy rule that was proposed by the U.S. Copyright Office in 2006 and was subsequently introduced, with some minor amendments, as a bill—the Orphan Works Act—in the U.S. Congress.⁸⁷ The rule would have limited the liability of those users who use an orphan work after an unsuccessful but reasonable search for the right owner has been conducted. It would not have eliminated the infringing nature of the use, but offered to good faith users of orphan works a closed set of remedies should the right owner reappear and start litigation over the use of the work. To qualify for this limited liability rule, the user would need to prove that he or she performed a “reasonably

83. Grimmelmann, *supra* note 29, at 13.

84. Samuelson, *supra* note 29, at 517 n.192.

85. *See supra* Section II.B.

86. *See* van Gompel, *supra* note 1, at 691–99. *See generally* VAN EECHOUDE ET AL., *supra* note 10, at 282–94 (reviewing the legislative models).

87. U.S. COPYRIGHT OFFICE (2006), *supra* note 3, at 71–89. The limitation on remedy rule found its way into the Orphan Works Act of 2006, H.R. 5439, 109th Cong. (2006) and the Copyright Modernization Act of 2006, H.R. 6052, 109th Cong. (2006). These bills died with the end of the Bush presidency. Later, two other bills were introduced, namely the Orphan Works Act of 2008, H.R. 5889, 110th Cong. (2008) and the Shawn Bentley Orphan Works Act of 2008, S. 2913, 110th Cong. (2008).

diligent search”⁸⁸ and, if possible and reasonably appropriate, to have provided attribution to the author or copyright owner of the work.⁸⁹

The proposed liability rule limits monetary relief to “reasonable compensation” for the use made. This should relate to a reasonable license fee as would have been established in negotiations between the user and right owner before the infringing use began.⁹⁰ However, if the use was non-commercial and the user expeditiously ceases the infringement upon a notice by the rights owner, no monetary relief is due at all.⁹¹ More importantly for transformative users is the limitation on injunctive relief. Where an orphan work has been incorporated into a derivative work, the copyright owner cannot obtain full injunctive relief to prevent the use of the derivative work, provided that the user pays a reasonable compensation and grants sufficient attribution to the right owner.⁹² Full injunctive relief remains available in cases where orphan works are used without any transformation of their content.

The proposed limitation on remedy rule is claimed to be a cost-efficient model in that it does not require users to compensate right owners in advance, but only in case they reappear and file a claim.⁹³ Still, the rule imposes several costs on users, including

- the costs of keeping records for being able to prove the diligence of search;
- the costs of assessing the likeliness of future claims; and

88. While the U.S. Copyright Office’s 2006 report, House Bill 5439, and House Bill 6052 speak of a “reasonably diligent search,” the House Bill 5889 and Senate Bill 2913 refer to a “qualifying search, in good faith” and define the criteria required for conducting such search. *See* U.S. COPYRIGHT OFFICE (2006), *supra* note 3, at 96–110; H.R. 5439; H.R. 6052; H.R. 5889; S. 2913. However, they do not describe them in minuscule detail, as search standards may well evolve with technology and will vary according to the creative sector and the nature of the use. *See* Jane C. Ginsburg, *Recent Developments in US Copyright Law: Part I – “Orphan” Works*, 217 REVUE INTERNATIONALE DU DROIT D’AUTEUR 99, 137 (2008).

89. The idea is that it should be unambiguously clear to the public that the work rightfully belongs to the author or copyright owner, and not to the user in question. *See* U.S. COPYRIGHT OFFICE (2006), *supra* note 3, at 110–12.

90. *See* Ginsburg, *supra* note 88, at 115–21 (considering reasonable compensation and what it includes). Note that House Bill 5889 allows courts to award extra compensation in case the orphan work has been registered. *See* H.R. 5889.

91. Under House Bill 5889, the beneficiaries of this “safe harbor” rule are non-profit educational institutions, libraries, archives and public broadcasting entities. *See* H.R. 5889. Senate Bill 2913 adds museums to this list. *See* S. 2913.

92. *See* Ginsburg, *supra* note 88, at 121–29 (criticizing the unavailability of injunctive relief for right holders whose works have been incorporated into a derivative work).

93. U.S. COPYRIGHT OFFICE (2006), *supra* note 3, at 115.

- the costs that arise in case a right holder reappears, including costs of litigation and of paying reasonable compensation after a successful litigation.⁹⁴

Moreover, it does not make effective use of the judicial system, because right holders are still required to file lawsuits against infringing users.⁹⁵ This casts doubts upon the actual cost efficiency of the rule.

Another question is whether the limitation on remedy rule would actually provide the legal certainty that transformative users require. This first depends on how courts would interpret the requirement of a “reasonably diligent search,” which is not well circumscribed in the proposals.⁹⁶ Even if that requirement were sufficiently clear, users would still face considerable difficulties if they had to convince a court *ex post* of the reasonableness of a search, especially where the search was conducted a long time ago. To provide evidence in court, they would need to exactly document each and every search they have made, and keep records thereof, possibly for an indefinite period of time. However, given the relatively small proportion of orphan works that would normally be incorporated into derivative works, transformative users would probably not perceive this as too big a problem to handle.

A second model, which would have advantages and disadvantages similar to those of the United States’ limitation on remedy rule, would be a specific exception or limitation for orphan works, as proposed by groups of

94. HUGENHOLTZ ET AL., *supra* note 1, at 190; Van Gompel, *supra* note 1, at 696; VAN EECHEUD ET AL., *supra* note 10, at 291.

95. Ginsburg, *supra* note 88, at 130–31. House Bill 5889 and Senate Bill 2913 state that before users can invoke the limitation on remedy rule in courts, they must first try to negotiate reasonable compensation in good faith with the copyright holder and, if agreement is reached, render payment of the compensation in a reasonably timely manner. *See* H.R. 5889; S. 2913. Negotiations of this kind appear to be aimed at keeping infringement claims out of the courts.

96. *See* Benjamin T. Hickman, *Can You Find a Home for This “Orphan” Copyright Work? A Statutory Solution for Copyright-Protected Works Whose Owners Cannot be Located*, 57 SYRACUSE L. REV. 123, 149 (2006) (criticizing the U.S. Copyright Office’s 2006 proposal “for not defin[ing] the steps a user would need to take to satisfy a reasonably diligent search”). Clarity on the meaning of a diligent search is also required to prevent works from being inaccurately labeled as orphans, thus ensuring that the legitimate interests of right holders are not unreasonably prejudiced. This should guarantee compliance of the rule with the three-step test of Art. 13 TRIPS Agreement. *See* Coree Thompson, Note, *Orphan Works, U.S. Copyright Law, and International Treaties: Reconciling Differences to Create a Brighter Future for Orphans Everywhere*, 23 ARIZ. J. INT’L & COMP. L. 787, 832–39 (2006); Ginsburg, *supra* note 88, at 139–41 (arguing that “the level of diligence should be set consistently high”).

stakeholders in the United Kingdom and Germany.⁹⁷ The orphan works exception would generally permit the use of an orphan work if despite reasonable search efforts⁹⁸ the right owner of the work could not be found. Once they reemerge, the right owners of works used under the exception are entitled by law to claim reasonable compensation for the use made.⁹⁹ The amount of compensation should either be agreed by negotiation, the failure of which shall result in the compensation being fixed by an independent third party, or set unilaterally by a CRMO.¹⁰⁰ Subsequent use of the work would be allowed only if the user succeeds in negotiating the terms of use with the right owner in the usual way. Where the work has been integrated or transformed into a derivative work, however, users should be allowed to continue using the work provided that a reasonable royalty is paid and sufficient acknowledgement is given to the right owner.¹⁰¹

A key difference with the U.S. liability rule is that, in the U.S. proposal, users of orphan works are deemed infringers, yet their liability is limited if they meet the statutory criteria. Under the orphan works exception, on the other hand, users of orphan works would not legally infringe any rights, provided that their use remains within the margins of the law. Moreover, resurfaced right holders do not have to sue for copyright infringement, as under the U.S. proposal, but can directly claim the compensation available under the exception. Right holders seeking compensation are thus released

97. See BRITISH SCREEN ADVISORY COUNCIL, *supra* note 6; Aktionsbündnis Urheberrecht, *supra* note 7. See also Gowers Review, *supra* note 7, at 69–72, Recommendation 13.

98. The BSAC proposal speaks of “best endeavours,” while the Aktionsbündnis Urheberrecht refers to “a reasonable professional and documented search” (“*einer angemessenen professionellen und dokumentierten Suche*”). BRITISH SCREEN ADVISORY COUNCIL, *supra* note 6; Aktionsbündnis Urheberrecht, *supra* note 7.

99. The reasonable compensation is called “reasonable royalty” in the BSAC proposal and “reasonable remuneration” (“*angemessene Vergütung*”) in the proposal of the Aktionsbündnis Urheberrecht. BRITISH SCREEN ADVISORY COUNCIL, *supra* note 6; Aktionsbündnis Urheberrecht, *supra* note 7.

100. The former is proposed by the BSAC, the latter by the Aktionsbündnis Urheberrecht. BRITISH SCREEN ADVISORY COUNCIL, *supra* note 6; Aktionsbündnis Urheberrecht, *supra* note 7.

101. Provision for subsequent usage of a previously orphaned work is only contained in the BSAC proposal. BRITISH SCREEN ADVISORY COUNCIL, *supra* note 6, at 33. The proposal of the Aktionsbündnis Urheberrecht contains no such rule. It generally states that right holders stepping forward cannot prevent the making available of their works by users who have satisfied the conditions set by law. The reason for this may perhaps be that the exception proposed by the Aktionsbündnis Urheberrecht is limited to the making available of orphan works (“*die öffentliche Zugänglichmachung von Werken, deren Urheber oder Rechteinhaber . . . nicht ermittelt werden können*”), thus excluding other uses such as the making of derivative works. See Aktionsbündnis Urheberrecht, *supra* note 7, at 1.

from initiating (possibly costly) court proceedings against users. Lawsuits would arise only if a right holder contests the reasonableness of a user's search, if a user fails to compensate a right holder for the use made, or if there is disagreement between right holders and users about the reasonableness of the compensation to be paid under the orphan works exception.¹⁰²

Other than this, the two models have comparable effects. Similar to the U.S. limitation on remedy rule, the orphan works exception would offer adequate legal certainty to users of orphan works only if the standards for a reasonable search are sufficiently clear.¹⁰³ In addition, as under the U.S. proposal, the correct application of the exception is tested only *ex post*, i.e., once the right holder has reappeared. This would require good faith users of orphan works to document their searches well and maintain adequate search records to prove the reasonableness of the search in case a right holder contests it and files a lawsuit against the user.¹⁰⁴

A third model that would provide optimal legal certainty to users of orphan works, yet is also the most costly in terms of establishment, maintenance, and operation, is the Canadian system that offers compulsory licenses to use orphan works. Pursuant to the Canadian Copyright Act, a user can apply to the Canadian Copyright Board to obtain a license to use a particular work in those cases where the identity or whereabouts of the rights owner cannot be ascertained by reasonable inquiry.¹⁰⁵ The Copyright Board must be satisfied that the applicant has made "reasonable efforts" to find the right owner before a license may be issued.¹⁰⁶ The purpose for which the

102. See VAN EECHOUDE ET AL., *supra* note 10, at 293.

103. Providing clarity on the criteria that define a reasonable search seems also necessary to ensure that an orphan works exception is in compliance with the three-step test laid down *inter alia* in Art. 13 TRIPS Agreement. See G. Spindler & J. Heckmann, *Retrodigitalisierung verwaister Printpublikationen: Die Nutzungsmöglichkeiten von "orphan works" de lege lata und ferenda*, *GRUR Int.* 271, 281–83 (2008).

104. See Van Gompel, *supra* note 1, at 699; VAN EECHOUDE ET AL., *supra* note 10, at 293.

105. Copyright Act, R.S.C. 1985, c. C-42, art. 77 (Can.). For comparable systems in other countries, see *supra* note 8.

106. It is not required that "every effort" has been made to trace the right holder, but an applicant must prove to have conducted a "thorough search." To that end, the applicant is advised to contact different CRMOs and publishing houses; to consult libraries, universities, and museums; to check the registers of copyright offices; to investigate inheritance records; and to simply search the Internet. See CANADIAN COPYRIGHT BOARD, UNLOCATABLE COPYRIGHT OWNERS (2001), available at <http://www.cb-cda.gc.ca/unlocatable-introuvables/brochure2-e.html>.

license is requested—e.g., commercial, educational, or religious—is irrelevant.¹⁰⁷

Once the Copyright Board is convinced that the applicant cannot locate the copyright owner despite reasonable efforts, it may grant a license regardless of whether the work is of domestic or foreign origin.¹⁰⁸ A license cannot be granted for works that are unpublished or works for which the publication status cannot be confirmed.¹⁰⁹ However, the Copyright Board has sometimes presumed prior publication if conclusive evidence was hard to come by, but the circumstances nevertheless indicated the likeliness of publication.¹¹⁰

The Copyright Board has full discretion to establish the appropriate terms and license fees in the circumstances of each particular case.¹¹¹ The license usually specifies the type of use that is authorized, the restrictions to this use, the expiry date, and other parameters. As a rule, it is non-exclusive and limited to the Canadian dominion. The Copyright Board cannot issue

107. LAURENT CARRIÈRE, UNLOCATABLE COPYRIGHT OWNERS: SOME COMMENTS ON THE LICENSING SCHEME OF SECTION 77 OF THE CANADIAN COPYRIGHT ACT § 5.2.4, at 9 (1998), available at <http://www.robic.ca/admin/pdf/277/103-LC.pdf>.

108. See, e.g., *Licence Application by the National Film Board of Canada for the Use of a Musical Work*, 2005-UO/TI-34 (Sept. 13, 2005) (application denied). Interestingly, in 21% of the applications filed, the Copyright Board manages to find the copyright owner before a license is issued. See Justice William J. Vancise, Chairman, Copyright Bd. of Can., Speech at a Seminar Jointly Sponsored by the Intellectual Property Institute of Canada and McGill University, at 5 (Aug. 15, 2007), available at <http://www.cb-cda.gc.ca/about-apropos/speeches-discours/20070815.pdf>.

109. Jerry Brito & Bridget Dooling, *An Orphan Works Affirmative Defense to Copyright Infringement Actions*, 12 MICH. TELECOMM. & TECH. L. REV. 75, 106 (2005), available at <http://www.mttr.org/voltwelve/brito&dooling.pdf>. The fact that the licensing scheme applies only to published works ensures that the moral right of the author to decide whether or not to make his or her work available to the public (i.e., the *droit de divulgation*) is respected. Yet, it can be viewed as a shortcoming of the system, as it is not always easy to resolve whether an old work or a photograph has ever been published. *But cf.* Peter B. Hirtle, *Unpublished Materials, New Technologies, and Copyright: Facilitating Scholarly Uses*, 49 J. COPYRIGHT SOC'Y U.S.A. 259, 265 (2001) (noting that the Internet has made it easier to discover if a work has previously been published).

110. See, e.g., Copyright Board Canada, File 1991-15, *Non-exclusive licence issued to U.P. Productions, Inc., Scarborough, Ontario, authorizing the reproduction of extracts from a published work* (Mar. 26, 1992) (accepting that the 1937 cartoon *Popeye: Lost and Foundry* has been published, although there was only circumstantial evidence relating to its distribution). *But see*, e.g., *Licence Application by the Canadian Centre for Architecture for the Reproduction of Photographs*, 2004-UO/TI-32 (Jan. 17, 2005) (application denied); *Licence Application by the Office of the Lieutenant Governor of Québec for the Reproduction of a Photograph*, 2004-UO/TI-37 (Mar. 3, 2005) (application denied).

111. See, e.g., *Licence Application by Breakthrough Films & Television for the Reproduction Through Off-Camera narration of Book Extracts in a Television Program*, 2004-UO/TI-33 (May 10, 2005). See also Vancise, *supra* note 108, at 6.

licenses beyond its own territory. The license also stipulates a royalty fee,¹¹² which often corresponds to an ordinary royalty rate as would have been made in consideration of consent being given.¹¹³

In most cases, the royalty fee is paid directly to the CRMO that would normally represent the untraceable right owner, but sometimes it must be deposited into an escrow account or trust fund. If the copyright owner surfaces, he may collect the royalties fixed in the license. If no right owner comes forward within five years after the expiry of the license,¹¹⁴ the collected royalty fee may be used for other purposes than those relating to the use in question.¹¹⁵

The Canadian model is often criticized for requiring an upfront payment of royalties that may never be distributed to the rightful copyright owner, due to the fact that they are unknown or untraceable.¹¹⁶ However, such disadvantage is the unintended result of a solution to a general market failure and therefore may need to be taken for granted. As in the ECL model, the non-distributable royalties could be used for cultural purposes or for establishing mechanisms that help to alleviate the problem of orphan works in the future. While this may imply that users end up paying royalties for means other than those relating to the use in question, it may be the price that a user needs to pay for the legal certainty to use an orphan work for which he or she would otherwise not be able to obtain the required permission.¹¹⁷

112. See Vancise, *supra* note 108, at 7 (giving reasons why the Copyright Board requires a royalty to be paid, even though the copyright owner is unknown and may perhaps never come forward).

113. See Carrière, *supra* note 107, at 9 (clarifying that the Copyright Board, in ascertaining the royalties to be paid, applies ordinary indicators such as “the tariffs of collectives for similar grant, comparison with similar licences granted outside the licensing schemes of section 77 or the circumstances and purposes of the requested licence”). There are cases involving charity-type uses, or situations where it is most likely that a work has entered the public domain, in which the Copyright Board requires the payment of royalties only if they are being claimed by a resurfaced copyright owner. See, e.g., *Non-Exclusive Licence Issued to the Canadian Institute for Historical Microreproductions Authorizing the Reproduction of 1,048 Works*, 1993-UO/TI-5 (Sept. 18, 1996). See also Vancise, *supra* note 108, at 7.

114. The statutory cut-off date to recover royalties is laid down in Copyright Act, R.S.C. 1985, c. C-42, art. 77(3) (Can.).

115. See CANADIAN COPYRIGHT BOARD, *supra* note 106 (explaining that the Copyright Board may for example instruct a CRMO to use the undistributed fees for the general benefit of its members).

116. See, e.g., De la Durantaye, *supra* note 78, ¶ 36.

117. See HUGENHOLTZ ET AL., *supra* note 1, at 194–95; Van Gompel, *supra* note 1, at 701.

Another more practical obstacle is that the legal infrastructure that would be required for a Canadian-style system is currently not available in all countries. There are many examples of countries that have no copyright board or tribunal that could clear the rights for orphan works.¹¹⁸ Nevertheless, a specialized court could always be entrusted with this task.¹¹⁹ A further strand of criticism extends to the operating costs of the Canadian model. The pre-clearance of orphan works by a public authority is said to be an expensive and lengthy process.¹²⁰ Although this may certainly hold true to a certain degree, it should not be exaggerated. The procedure to obtain a license at the Copyright Board is free of charge. The Canadian Copyright Board moreover indicates that, once it has received all the required information, a decision can usually be issued within thirty to forty-five days.¹²¹

The key advantage of the Canadian model is that it offers adequate legal certainty to users of orphan works. A license granted by the Copyright Board authorizes the user to use an orphan work without the risk of an infringement claim should the right owner come forward. As the type of uses for which the Canadian Copyright Board can give authorization includes the making of adaptations,¹²² obtaining a license would grant optimal relief to

118. See P.B. HUGENHOLTZ, ET AL., MINISTERIE VAN JUSTITIE, GESCHILLENBESLECHTING EN COLLECTIEF RECHTENBEHEER: OVER TARIEVEN, TRANSPARANTIE EN TRIBUNALEN IN HET AUTEURSRECHT [DISPUTE RESOLUTION AND COLLECTIVE MANAGEMENT OF RIGHTS: ON TARIFFS, TRANSPARENCY AND TRIBUNALS IN COPYRIGHT] (2007), available at http://www.ivir.nl/publicaties/hughenoltz/WODC_geschillenbeslechting_2007.pdf (examining the feasibility and desirability of creating a copyright tribunal in the Netherlands).

119. Van Gompel, *supra* note 1, at 694.

120. See, e.g., Jerry Brito & Bridget Dooling, *supra* note 109, at 106–07; Dennis W.K. Khong, *Orphan Works, Abandonware and the Missing Market for Copyrighted Goods*, 15 INT'L J.L. & INFO. TECH. 54, 75 (2007). The opponents of the Canadian model also maintain that the inefficiency of the system is exposed by the small number of applications filed before the Board. Indeed, from 1989 through August 2012, the Copyright Board has issued only 264 licenses out of 272 applications that were filed. See *Unlocatable Copyright Owners: Decisions/Licences Issued*, COPYRIGHT BOARD OF CANADA, <http://www.cb-cda.gc.ca/unlocatable-introuvables/licences-e.html> (last visited Aug. 29, 2012). The relatively small number of applications, however, might also be due to other factors, for instance, by the inability of the Copyright Board to grant licenses other than for uses within Canada.

121. CANADIAN COPYRIGHT BOARD, *supra* note 106.

122. Copyright Act, R.S.C. 1985, c. C-42, art. 77(1) (Can.) allows licenses to be issued for uses specified in arts. 3, 15, 18 and 21. This includes the making of adaptations. See JEREMY DE BEER & MARIO BOUCHARD, CANADA'S "ORPHAN WORKS" REGIME: UNLOCATABLE COPYRIGHT OWNERS AND THE COPYRIGHT BOARD 11 (2009), available at <http://www.cb-cda.gc.ca/about-apropos/2010-11-19-newstudy.pdf> (explaining that Art. 3 Canadian Copyright Act covers "the right to produce, reproduce, perform, translate, convert, record, adapt, telecommunicate and exhibit a work in public").

users who wish to incorporate an orphan work into a derivative work. At the same time, the system does not unnecessarily compromise the legitimate interests of the right owners concerned.¹²³ First, the verification of the good faith of a user being performed by an independent public body serves the needs of prospective users while taking due account of right owners' concerns. Second, decisions to grant a license are made on a case-by-case basis, thereby avoiding disproportional prejudice to the principle of exclusive rights that is inherent to the copyright system. Third, the license is granted to a particular user for a specific kind of use only. Fourth, the system does not result in a loss of income for right holders. If a right holder resurfaces, he is compensated.

C. RESTRAINING THE GOAT: PRACTICAL LEGAL SOLUTIONS FOR SMALL-SCALE INCIDENTAL USES

The last category of uses for which it must be determined what would be the best approach to deal with orphan works is small-scale incidental uses. As observed in Section II.C, *supra*, in Europe, the problem of orphan works has a relatively small impact of these types of uses. This can in part be explained by the absence of statutory damages with a punitive nature in Europe and the fact that such uses are not always threatened by injunctions. Furthermore, orphan works certainly are not a “new” problem.¹²⁴ For centuries, users in Europe have coped with the difficulties in tracing the right owners of works.¹²⁵ Rather than not use these works, however, they have found practical ways to handle this situation.

First, in cases of small-scale incidental use of an orphan work that are not really harmed by a possible injunction, good faith users in Europe sometimes post a legal disclaimer, indicating that in good faith they used material of which they were unable to locate the right owner and ask for permission. An example from the Netherlands is the following disclaimer that features prominently in a CD booklet of Wim Sonneveld, a famous Dutch artist of the 1950–1970's:

123. Moreover, the Canadian model appears compatible with the substantive minima of the Berne Convention, e.g., with the prohibition on formalities of Art. 5(2). *See* 1 Sam Ricketson & Jane C. Ginsburg, *INTERNATIONAL COPYRIGHT AND NEIGHBOURING RIGHTS: THE BERNE CONVENTION AND BEYOND* 329 (Oxford University Press 2d ed. 2006).

124. *See* HUGENHOLTZ ET AL., *supra* note 1, at 162.

125. *See* P. Bernt Hugenholtz & Annemique M.E. de Kroon, *The Electronic Rights War: Who Owns the Rights to New Digital Uses of Existing Works of authorship?*, 4 *IRIS* 16, 16 (2000) (explaining that technological developments have always sparked new secondary uses of pre-existing works).

Despite our best efforts, we were unable to find all right holder of the photographic works used in this booklet. If you believe, in good faith, to have a legitimate copyright claim, we kindly invite you to contact [X].¹²⁶

Likewise, in the United Kingdom, a recent report on managing orphan works suggests good faith users to include with every use of the orphan work a prominently visible disclaimer, stating:

Every effort has been made to trace copyright holders, obtain permission from them and to ensure that all credits are correct. [X] has acted in good faith at all times and on the best information available to us at the time of publication. We apologise for any inadvertent omissions, which will be corrected as soon as possible if notification is given to us in writing.¹²⁷

Obviously, disclaimers of this kind provide no immunity against liability. However, they may help to demonstrate the good faith intention of the user. After all, by adding a disclaimer, the user makes no secret of having used an orphan work. Moreover, by inviting the right holder, should he or she emerge, to contact the user to arrange a deal, the disclaimer attempts to resolve possible conflicts with reappearing right holders through an amicable settlement. In Europe, therefore, the use of disclaimers to indicate good faith uses of orphan works is totally accepted. Although not widely applied, the disclaimers appear in many different contexts.¹²⁸

A second practical solution that has emerged in Europe to address the problem of orphan works is the grant of an indemnity or security by a CRMO or a private organization representing right holders to prospective users of orphan works. CRMOs sometimes include an indemnity clause in collective licensing agreements with users, in which they assume the financial liability for any claim made by a copyright owner who is not represented by them.¹²⁹

126. See Wim Sonneveld, *Voor altijd* (Nikkelen Nelis 1999). The disclaimer contained in the accompanying booklet states: “*Ondanks onze pogingen zijn we er niet in geslaagd alle rechthebbenden op het gebruikte fotomateriaal te achterhalen. Diegenen die menen alsnog aanspraken te kunnen doen gelden, worden verzocht contact op te nemen met [X].*”

127. NAOMI KORN & EMMA BEER, JOINT INFO. SYS. COMM., BRIEFING PAPER ON MANAGING ORPHAN WORKS 3 (2011), available at <http://www.jisc.ac.uk/media/documents/publications/programme/2011/scaorphanworksbp.pdf>.

128. They appear inter alia in the closing credits of television productions and sometimes on websites. See, e.g., the disclaimer used by the Dutch broadcaster NPS, *supra* note 52; *Disclaimer*, FILMEDUCATIE.NL, <http://www.filmeducatie.nl/disclaimer> (last visited Aug. 29, 2012).

129. Koskinen-Olsson, *supra* note 66, at 292.

Outside the field of collective licensing there are also instances of indemnities being applied. In some countries, indemnities or securities are issued by organizations representing a specific category of right holders. In the Netherlands, for example, users who fail to identify or locate the right holder of a photographic work that they wish to reutilize can request *Foto Anoniem*,¹³⁰ a foundation linked to Burafo, a Dutch organization for professional photographers, to grant them a contractual indemnity. The indemnity exempts users from liability for financial damages. It is issued upon payment of a fair compensation. The compensation, which is similar to the license fee that would normally be charged for publication of a comparable photograph, is reserved by *Foto Anoniem* to pay rights holders in case they subsequently come forward. *Foto Anoniem* actively searches for right holders of photographs for which it has issued an indemnity and keeps an online record of indemnified photographs.¹³¹

A contractual indemnity or security protects against financial liability for copyright infringement, thus providing a measure of legal certainty to users of orphan works who have obtained an indemnity or security. Nevertheless, it does not offer a complete solution to the orphan works problem. The fact that an indemnity or security has been granted does not prevent reappearing right holders from seeking injunctive relief. Hence, the further use of the works could still be prohibited.¹³²

These two examples demonstrate that in Europe users have found practical ways to deal with the orphan works problem for particular small-scale incidental types of uses. Because of the availability of statutory damages in U.S. copyright law,¹³³ on the other hand, it seems virtually impossible to adopt similar approaches in the United States. First, using a disclaimer of the kind applied in Europe would be highly risky, as it could possibly be held against the user in a copyright infringement proceeding.¹³⁴ That is, if a user

130. FOTO ANONIEM [ANONYMOUS PHOTO], <http://www.fotoanoniem.nl> (last visited Aug. 29, 2012).

131. In Belgium, a similar model is employed by SOFAM, the CRMO for visual arts, which upon request grants suretyship (*borgstellingen*) to users who are unable to trace the copyright owner. See Contracten, SOFAM, <http://www.sofam.be/nl/10/> (last visited Aug. 29, 2012).

132. Van Gompel, *supra* note 1, at 690.

133. See *supra* Section II.A.

134. At least it seems to bar an innocent intent defense in mitigation of damages. Pursuant to 17 U.S.C. § 504(c)(2) (2010), courts can reduce the award of statutory damages if the infringer “was not aware and had no reason to believe that his or her acts constituted an infringement of copyright” or shall remit statutory damages if the infringer is a nonprofit educational institution, library, or archive that “believed and had reasonable grounds for believing that his or her use of the copyrighted work was a fair use under section 107.” By

would openly disclose that despite reasonable efforts it was impossible to find the right owner to ask for permission, this could perhaps be used as evidence to prove that the user had constructive knowledge of the infringing nature of his or her conduct. This could give rise to the risk of liability for payment of statutory damages for willful infringement.¹³⁵ For the same reason, granting indemnities or securities to good faith users of orphan works would probably be perceived as an equally unattractive model. Because of the risk of exposure to statutory damages, it may be difficult to find CRMOs or private organizations representing right holders in the United States to provide financial surety for users of orphan works.

Thus, users who wish to make incidental use of an orphan work in the United States cannot effectively rely on practical models existing in Europe. Addressing the orphan works problem for this group of users requires a legislative approach. Given the small scale of use, the models based on a diligent search requirement as discussed in Section III.B, *supra*, could well provide relief to these users. In seeking a model to facilitate the use of orphan works in the United States, therefore, this category of users could easily be grouped together with transformative users.

IV. CONCLUSION

The orphan works problem is multifaceted. It manifests itself differently according to how the works are used. This article has metaphorically represented the problem of orphan works and its different manifestations by the image of the Chimera. The Chimera is a three-headed she-monster, which according to ancient Greek mythology had a lion's head, a tail that

using a disclaimer stating that despite reasonable efforts it was impossible to find the copyright owners of the works used, the user shows to have consciously and intentionally used works without permission, thus confirming knowledge of the infringing nature of his or her conduct. Also, for the nonprofit educational institution, library, or archive relying on fair use, the use of such a disclaimer would not make much sense, as it can be regarded as an invitation to the right owner to arrange an ex post licensing deal, thus negating to a greater or lesser degree the reliance on fair use. Accordingly, even though the innocent intent defense in mitigation of damages is hardly ever accepted by the courts, *see* Samuelson & Wheatland, *supra* note 42, at 452–54, 474, it would still be better to omit disclaimers of this kind if one were to rely on this provision in litigation.

135. 17 U.S.C. § 504(c) (2010). Whether statutory damages for willful infringement are actually awarded will of course depend on the judicial interpretation of willfulness. However, some scholars have observed that “[c]ourts have interpreted willfulness expansively such that infringement may be deemed willful if the defendant should have known his conduct was infringing. Judges have even found infringement to be willful as to defendants who proffered plausible, if ultimately unsuccessful, fair use defenses.” Samuelson & Wheatland, *supra* note 42, at 459. As a consequence, “virtually all ordinary infringers are at risk of excessive statutory damage awards.” *Id.* at 455.

ended in a serpent's head, and a head of a goat in her middle. The three animal parts symbolize the three most characteristic appearances of the problem.

For mass-digitization projects (the lion's head of the Chimera), the orphan works problem is particularly urgent because of the large scale in which pre-existing works are being digitized and made available online. These projects face considerable uncertainty when reusing works where the right owners are unknown or untraceable. Because of the large scale of use, the risk of becoming liable for payment of monetary damages is particularly high for mass-digitization projects. By contrast, these projects fear relatively little from injunction claims.

This is different for users who incorporate orphan works into a derivative work (the serpent's head of the Chimera). The insecurity that the orphan works problem causes for these users is that a resurfaced right owner may take legal action and claim injunctive relief by prohibiting further use of the orphan work. This could seriously harm the derivative work. On the other hand, the risk of incurring monetary damages seems bearable only if a few orphan works are integrated into the derivative work. Still, the availability of statutory damages may render the financial risk of using orphan works too high for certain creative re-users in the United States.

A third category is certain small-scale incidental uses of orphan works (the goat's head of the Chimera). Because in Europe no statutory damages apply and small-scale incidental uses are not easily threatened by injunctions, the orphan works problem is not perceived by users in this group as being too serious an issue. In certain European countries, voluntary practices have also been developed to deal with orphan works. These practices, which exist in the use of disclaimers and indemnities, have further decreased the significance of the issue for small-scale users in Europe. In the United States, on the other hand, statutory damages effectively deter users from relying on practical models. In contrast to Europe, addressing the orphan works problem in the United States would therefore require taking legislative action for these types of uses too.

The fact that the problem of orphan works manifests itself in different ways indicates that, in order to address it effectively, a multifaceted strategy ought to be pursued. Other than relying on practical legal solutions that are already in place in certain European countries (but could perhaps be voluntarily extended to other fields of copyright law where such mechanisms do not yet exist), the strategy proposed here is to take legal action on two fronts.

First, for mass-digitization projects, a model based on collective licensing appears to be most adequate and effective. Collective licensing by itself

would already provide effective relief in cases where permission to use a multitude of works needs to be obtained. Should an extended collective licensing model be introduced, however, this would give greater legal certainty. It would ensure that non-represented right holders, including right holders of orphan works, are also covered by the licensing scheme. While this model has the advantage of steering close to the tradition of collective rights management that has a proven track record, particularly in Europe, it also comes with a number of disadvantages and practical difficulties, one of which is the upfront payment of royalties that it requires. Nonetheless, as observed, such cost would not outweigh the benefit of not having to clear rights on a work-by-work basis.

Second, for users who incorporate an orphan work into a derivative work (and in the United States also for small-scale incidental users of orphan works), a legislative model based on diligent search requirements could be adopted. Examples include the U.S.-style limitation on remedy rule, a specific orphan works exception, or the Canadian model of compulsory licensing by a public authority. All three models have specific advantages and disadvantages. The Canadian model provides maximum legal certainty, but comes with upfront costs. A limitation on remedy rule or orphan works exception offers lesser legal certainty to users of orphan works, but does not require advance payments (though surely also entail *ex ante* costs). Which model is the most suitable, therefore, will vary according to the specific national situation.

In short, by pursuing the multifaceted strategy proposed in this Article, the problem of orphan works can most adequately be addressed for the different groups of users that are confronted with it. As a practical approach sometimes suffices, at least in Europe, legislative intervention is only required in cases where there is true legal uncertainty for users. Hence the strategy put forward here is not aimed at entirely eliminating the Chimera, but at taming the lion, killing the serpent, and restraining the goat. Only in the United States must the goat be killed together with the serpent. This should prevent small-scale incidental users of orphan works from exposure to the risk of becoming liable for payment of (perhaps exorbitant) statutory damages.

HOW FAIR USE CAN HELP SOLVE THE ORPHAN WORKS PROBLEM

Jennifer M. Urban[†]

ABSTRACT

Many works that libraries, archives, and historical societies would like to digitize are “orphan works,” that is, works for which the copyright holder either is unknown or cannot be located after a diligent search. Due to copyright risk if an owner later shows up, nonprofit libraries and similar institutions have been reluctant to digitize and make these works available, greatly limiting access to important cultural and historical information. While a legislative fix may soon be proposed, this Article argues that legislation is not necessary to enable some uses of orphan works by nonprofit libraries and archives. Instead, U.S. copyright law’s fair use doctrine, which allows certain unpermitted uses of copyrighted works, provides a partial solution. Though it is an incomplete solution, fair use has some significant advantages over other approaches through which libraries and archives could make publicly beneficial uses of orphan works. Under fair use, there is no need to develop a licensing system, significantly reducing administrative and transactional costs, and eliminating socially wasteful license fees for works that are not on the market, and for which an owner is unlikely to exist. Second, fair use has the flexibility to accommodate change over time as libraries and archives discover the best ways to search for owners, preserve works, and make them available. Finally, allowing fair use of orphans by libraries and archives helps fulfill copyright’s critical purposes of promoting the dissemination of knowledge and supporting speech and free expression.

© 2012 Jennifer M. Urban.

[†] Assistant Clinical Professor of Law, University of California, Berkeley School of Law. The author gratefully acknowledges the many thoughtful ideas and helpful comments of Ty Alper, Roxanna Altholz, Michelle Anderson, Michael Donaldson, James Grimmelman, David Hansen, Edward Lee, Mark Lemley, Saira Mohammed, Pamela Samuelson, Jason Schultz, Christopher Sprigman, and Karen Tani, participants in the Orphan Works & Mass Digitization: Obstacles & Opportunities conference [hereinafter Berkeley Orphan Works Conference], and participants in the 12th Annual Intellectual Property Scholars Conference. This Article is output of the Berkeley Digital Library Copyright Project, generously supported by funding from the Alfred P. Sloan Foundation. More information about the Berkeley Digital Library Project can be found at <http://www.law.berkeley.edu/12040.htm>. I especially thank Ana Enriquez for excellent research assistance. The thoughts contained in this Article represent my individual academic conclusions and opinions, and should not be attributed to any client I represent or the Samuelson Law, Technology & Public Policy Clinic, which I direct. In addition, the University of California, where I am a faculty member, is a defendant in the *Authors Guild, Inc. v. HathiTrust* suit; however, I do not represent, and in no way speak for, the University of California or any other party in the case.

TABLE OF CONTENTS

I.	INTRODUCTION.....	1380
II.	BRIEF BACKGROUND ON ORPHAN WORKS AND FAIR USE.....	1387
III.	THE ORPHAN STATUS OF WORKS IS HIGHLY RELEVANT TO THE NATURE-OF-WORK AND MARKET-HARM FACTORS AND FAVORS FAIR USE.....	1392
	A. THE “NATURE” OF AN ORPHAN WORK	1392
	1. <i>Orphan Status as a Key Aspect of a Work’s “Nature”</i>	1395
	2. <i>Other Characteristics of Works in Library and Archive Collections</i>	1401
	3. <i>The Market-Harm Factor and Orphan Works</i>	1402
	B. MARKET HARM	1404
	C. COLLECTIVE OR OTHER CENTRALLY MANAGED LICENSING	1409
IV.	NONCOMMERCIAL USE OF AND ACCESS TO ORPHAN WORKS VIA LIBRARIES AND ARCHIVES PROMOTES THE PROGRESS OF SCIENCE.....	1414
	A. THE PURPOSE AND CHARACTER OF NONPROFIT LIBRARY AND ARCHIVE USES OF ORPHANS FAVORS FAIR USE.....	1414
	B. LIBRARY AND ARCHIVE DIGITIZATION OF ENTIRE ORPHAN WORKS OFTEN SHOULD NOT WEIGH AGAINST FAIR USE UNDER THE AMOUNT-AND-SUBSTANTIALITY FACTOR	1422
V.	FURTHER STRENGTHENING FAIR USE’S APPLICATION	1425
VI.	CONCLUSION	1427

I. INTRODUCTION

The twentieth century witnessed an explosion in creative production. By 1952, approximately 250,000 books were being published a year; by the year 2000, it was one million.¹ At the same time, new forms of media—professional and amateur movies and photography, relatively inexpensive print media, and microforms, to name a few—developed and were used to create countless works, ranging from high art, to family photographs, to quotidian organizational records. Today, most such materials are only

1. GABRIEL ZAID, SO MANY BOOKS: READING AND PUBLISHING IN AN AGE OF ABUNDANCE 20–21 (1st ed. Paul Dry Books 2003).

available to the public via the limited reach of physical collections. However, advances in digitization technologies offer tantalizing possibilities for preserving and providing greater access to our cultural and scientific heritage. Online access to digital copies of library and archive collections holds the promise of connecting many more readers with research, historical materials, and cultural ephemera.

Some digitization projects currently being undertaken by entities large and small illustrate the potential benefits of digitization programs. For example, the Library of Congress (LOC) makes millions of items available online through its American Memory project, a digitized collection of “written and spoken words, sound recordings, still and moving images, prints, maps, and sheet music that document the American experience.”² The Prelinger Archive collects ephemeral and home movies, and makes many of them available through the Internet Archive.³ The Civil Rights Movement Veterans website collects, digitizes, and puts online original documentation of the civil rights movement, from photographs to the organizational meeting minutes of the Student Nonviolent Coordinating Committee.⁴ A variety of other projects exist as well.⁵ Many of the works in these collections are rare, obscure, or simply specialized in nature. These works are unlikely to be commercially valuable or to circulate widely from traditional print

2. *Mission and History (American Memory)*, LIBRARY OF CONGRESS, <http://memory.loc.gov/ammem/about/index.html> (last visited July 28, 2012) [hereinafter LIBRARY OF CONGRESS, *Mission*]; see, e.g., *An American Time Capsule: Three Centuries of Broad-sides and Other Printed Ephemera*, LIBRARY OF CONGRESS, <http://memory.loc.gov/ammem/rbpehtml/> (last visited July 28, 2012).

3. *Welcome to Prelinger Archives*, INTERNET ARCHIVE, <http://archive.org/details/prelinger> (last visited July 28, 2012).

4. *Documents of the Southern Freedom Movement 1951–1968*, CIVIL RIGHTS MOVEMENT VETERANS, <http://www.crmvet.org/docs/dochome.htm> (last visited July 28, 2012); *Photo Album*, CIVIL RIGHTS MOVEMENT VETERANS, http://www.crmvet.org/images/img_home.htm (last visited July 28, 2012). See also Bruce Hartford, Webmaster, Civil Rights Movement Veterans Website, Panel at Berkeley Orphan Works Conference: Who Wants to Make Use of Orphan Works and Why? (Apr. 12, 2012) (slides available at <http://www.law.berkeley.edu/files/Hartford.pdf>).

5. See, e.g., *Patriotic Melodies*, LIBRARY OF CONGRESS, <http://lcweb2.loc.gov/diglib/ahas/html/patriotic/patriotic-home.html> (last visited July 24, 2012) (one of many such projects at the Library of Congress); *Michigan Digitization Project*, MICHIGAN LIBRARY, <http://www.lib.umich.edu/michigan-digitization-project> (last visited July 24, 2012); Allie Shay, *Digitizing the Long Civil Rights Movement in North Carolina*, PUBLISHING THE LONG CIVIL RIGHTS MOVEMENT (Mar. 13, 2012), <https://lcrm.lib.unc.edu/blog/index.php/2012/03/13/digitizing-the-long-civil-rights-movement-in-north-carolina>.

collections,⁶ but in the LOC's words, they can embody "a digital record of American history and creativity," and ultimately, "the 'nation's memory.'" ⁷

While works like these collectively comprise "the nation's memory," a significant proportion of such materials are also likely to be "orphan works"—presumptively copyrighted works for which no owner can be located. In such cases, there is no feasible way for libraries and archives to obtain permission to digitize and make the works available. By digitizing orphan works and putting them online, libraries and archives thus take the risk of a putative owner later appearing and suing. While the LOC and others have moved ahead with the digitization of some orphan works, the possibility of damages or injunctive relief still chills a great deal of potentially-socially beneficial activity.⁸ For this reason among others, in 2006 the Copyright Office recommended legislation as a way to free up reuses of orphan works. While this legislation ultimately did not pass, the current Register of Copyrights, Maria Pallante, recently announced plans to pursue new orphan works legislation in Congress in the near future.⁹

6. See ASS'N OF RES. LIBR., ARL STATISTICS 2008–2009, at 9 (2011), available at <http://publications.arl.org/1h8eiv.pdf> (showing low and decreasing circulation of print collections among research libraries). Decreasing circulation trends are even more pronounced when looking at individual collections. A study of the Cornell University Library, for example, shows that approximately fifty-five percent of its print monographs acquired since 1990 have never actually circulated. Cornell Univ. Libr., Report of the Collection Development Executive Committee Task Force on Print Collection Usage Cornell University Library 2, 13 (2010), available at http://staffweb.library.cornell.edu/system/files/CollectionUsageTF_ReportFinal11-22-10.pdf; see also Nicholas Joint, *Is Digitisation the New Circulation?: Borrowing Trends, Digitisation and the Nature Of Reading in US and UK Libraries*, 57 LIBR. REV. 87, 90–92 (2008) (describing "digital natives[']" preference for access to digital versus print collections).

7. LIBRARY OF CONGRESS, *Mission*, *supra* note 2.

8. U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS: A REPORT OF THE REGISTER OF COPYRIGHTS 15 (2006).

Many users of copyrighted works who have limited resources or are particularly risk-averse have indicated that the risk of liability for copyright infringement, however remote, is enough to prompt them simply to not make use of the work. Such an outcome is not in the public interest, particularly where the copyright owner is not locatable because he no longer exists or otherwise does not care to restrain the use of his work.

Id.

9. Maria A. Pallante, *Orphan Works & Mass Digitization: Obstacles & Opportunities*, 27 BERKELEY TECH. L.J. 1251 (2012). As this Article went to press, the Copyright Office issued an initial Notice of Inquiry on the orphan works problem, confirming that the Office is undertaking a new process to delve into the problem. Notice of Inquiry: Orphan Works and Mass Digitization, 77 Fed. Reg. 64,555 (Oct. 22, 2012).

This Article, however, explains why legislation should not be necessary to enable certain kinds of uses of orphan works. Specifically, it argues that courts should apply U.S. copyright law's fair use doctrine to allow the LOC and other nonprofit libraries and archives¹⁰ to digitize appropriately identified orphan works¹¹ in their collections, and to publicly display these works to their patrons for noncommercial acts such as reading, viewing, conducting research, and developing scholarship.¹² Some institutions have already been willing to move forward. It appears that the LOC, for example, has been relying on the fair use doctrine as a justification for digitizing and putting orphan works in its American Memory collections online for noncommercial use.¹³ Other libraries and archives have also made some orphan works in their collections available, despite the impossibility of clearing them.¹⁴ These

10. This is not to suggest that other types of entities or individuals are ineligible to claim fair use for digitizing orphans; this Article simply does not take up this question.

11. Orphan works, of course, must be properly identified as such. The Copyright Office Report identified the need for a reasonably "diligent search" to identify orphans, and the subsequent legislation introduced in Congress tracked that recommendation. U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS, *supra* note 8, at 96; Shawn Bentley Orphan Works Act of 2008, S. 2913, 110th Cong. (2008); Orphan Works Act of 2008, H.R. 5889, 110th Cong. (2008); Orphan Works Act of 2006, H.R. 5439, 109th Cong. (2006). For purposes of this Article, I am leaving aside the details of what might constitute a reasonably diligent search and simply assuming that one has been done by the library or archive wishing to digitize a work, and that no owner was locatable. *See infra* note 61 for a further brief discussion.

12. Note that this Article takes no position about whether fair use is an appropriate solution for uses of orphan works outside of the nonprofit library and archival uses described above. For example, this Article does not consider making derivative works from orphan works, nor does it consider commercial uses at all. For some such uses, fair use may also appropriately play a role. For others, other proposed solutions—such as the original proposal by the U.S. Copyright Office to limit remedies for uses made after a reasonable search for the copyright owner—may instead be most appropriate.

13. *See, e.g., American Women: Copyright and Other Restrictions*, LIBRARY OF CONGRESS, <http://memory.loc.gov/ammem/awhtml/awcopy.html> (last visited July 24, 2012) ("In [some] cases, the Library was unable to identify a possible rights holder and has elected to place some of those items online as an exercise of fair use for strictly non-commercial educational uses."); *The Wilbur and Orville Wright Papers: Copyright and Other Restrictions*, LIBRARY OF CONGRESS, <http://memory.loc.gov/ammem/wrighthtml/wrightres.html> (last visited July 24, 2012).

14. At the Berkeley Orphan Works Conference, for example, speakers from archives offered examples of a wide range of works they are making available in digital form. Brewster Kahle, Founder, Internet Archive; Kenneth Crews, Director, Columbia Copyright Advisory Office; Bruce Hartford, Webmaster, Civil Rights Movement Veterans Website; & Rick Prelinger, Founder, Prelinger Archives, Panel at Berkeley Orphan Works Conference: Who Wants to Make Use of Orphan Works and Why? (Apr. 12, 2012) (slides available at <http://www.law.berkeley.edu/11731.htm>); *see also* Shay, *supra* note 5; *About the Grateful Dead Archive*, UNIVERSITY OF CALIFORNIA, SANTA CRUZ UNIVERSITY LIBRARY, <http://library.ucsc.edu/grateful-dead-archive/about> (last visited July 24, 2012).

activities do much to further the constitutional value of “promot[ing] the Progress of Science,”¹⁵ by freeing important features of the “nation’s memory” and placing them into the hands of researchers, students, and other members of the citizenry, where they can illuminate our cultural history and support the further development of knowledge.

Indeed, given the lack of harm to owners who are unlocatable—who, indeed, may not exist at all—and the social benefits that would result from libraries and archives digitizing orphans and making them accessible, the conclusion that fair use covers such activities may seem obvious. This conclusion, however, is not accepted by all. The recent lawsuit brought by the Authors Guild against the HathiTrust illustrates this point.¹⁶ The case involved a “large scale digital preservation repository”¹⁷ of nearly ten million books, some of which are orphans, run by a consortium of more than sixty research libraries and some of its member libraries. Among claims against various aspects of the HathiTrust’s activities, the Authors Guild specifically argued that fair use does not justify the HathiTrust’s digitization and access program for orphans.¹⁸ Additionally, libraries and archives themselves have expressed concerns about the perceived uncertainty of relying on fair use in order to provide access to orphans.¹⁹ As this Article was going to press, the

15. U.S. CONST., art. I, § 8, cl. 8.

16. As this Article was going to press, the Southern District of New York released its decision in *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939 (S.D.N.Y. Oct. 10, 2012), an important fair use case considering the mass digitization of (non-orphans) works. The Article has been updated to reflect relevant aspects of this decision.

17. *Guidelines for Digital Object Deposit*, HATHITRUST DIGITAL LIBRARY, http://www.hathitrust.org/deposit_guidelines (last updated Dec. 10, 2011).

18. Memorandum of Law in Support of Plaintiffs’ Motion for Partial Judgment on the Pleadings at 21–23, *Authors Guild, Inc. v. HathiTrust*, No. 11 Civ. 6351 (S.D.N.Y. Feb. 28, 2012). Note that the plaintiffs did not analyze fair use in detail in their motion for partial judgment on the pleadings, but instead argued that fair use can never apply to the defendants because it falls outside of the boundaries of § 108 of the Copyright Act’s specific exceptions for libraries. This argument seems fundamentally misguided; it appears to be incorrect in light of the plain language of § 108 as well as the purposes animating both copyright law generally and fair use specifically. See Jonathan Band, *The Impact of Substantial Compliance with Copyright Exceptions on Fair Use* 3–5 (Nov. 30, 2011) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1966593. And in fact, as this Article was going to press, Judge Baer’s decision in the *HathiTrust* case decisively dismissed it, stating, “Here, fair use does not undermine § 108, but rather supplements it.” *Authors Guild*, 2012 WL 4808939, at *8–9.

19. See, e.g., PRUDENCE ADLER, BRANDON BUTLER, PATRICIA AUFDERHEIDE & PETER JASZI, *FAIR USE CHALLENGES IN ACADEMIC AND RESEARCH LIBRARIES* 6, 13–15 (2010), available at http://www.arl.org/bm~doc/arl_csm_fairusereport.pdf; Marc Parry, *Out of Fear, Colleges Lock Books and Images Away from Scholars*, CHRONICLE HIGHER ED. (May 29, 2011), <http://chronicle.com/article/Out-of-Fear-Institutions-Lock/127701/>.

HathiTrust case was decided in the trial court.²⁰ Though the court did not reach the question of orphan works specifically,²¹ Judge Harold Baer decisively held that fair use protected several uses of works in the corpus by known authors made by the HathiTrust libraries. He also held that the digitization required to make those uses possible was fair.²²

This Article demonstrates that courts should also consider it fair use for libraries and archives to make nonprofit uses of orphan works, including a broader array of uses than the court considered for the non-orphan works in the *HathiTrust* case. At its core, the Article posits three questions: First, should reproducing (by digitization or other means) an orphan work by a nonprofit library or archive for noncommercial, socially beneficial purposes—for example, preservation, education, access to knowledge, traditional forms of research and scholarship, and new forms such as text mining—be considered a fair use? Second, should making the copy available to patrons—for example, by posting it online for viewing—also be considered a fair use? More generally, the Article considers whether the fair use doctrine should do any additional work when the copyrighted work in question is an orphan.

I answer each of these questions in the affirmative. Although there is no directly analogous case on point²³ and some aspects of the use—for example,

20. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939 (S.D.N.Y. Oct. 10, 2012).

21. *Id.* at *7–8.

22. *Id.* at *9. These uses include digitization of the works for the purposes of search, preservation, and accessibility for patrons with print disabilities. *Id.* at *2 (explaining the uses covered by the decision). These uses closely track some of the uses that I consider in Section IV.A, *infra*, and analyze as likely to be fair uses of orphans. As the first case to consider these uses by libraries, Judge Baer’s decision is supportive of these aspects of my argument; as discussed further throughout the Article, the reasons why these uses are fair should apply with even greater force for orphans. It should be noted that the libraries’ Orphan Works Program (OWP), which would have provided full-text access to works to library patrons, rather than the bare search result provided for all works, was not ripe to be adjudicated. *Id.* (describing the full-text access for patrons); *id.* at *7–8 (dismissing claims at OWP as not ripe for adjudication). As such, the case does not decide whether the claim made in this Article—that full-text access of true orphans should be considered fair use—is correct.

23. The *Authors Guild, Inc. v. HathiTrust* case may confront this question directly, but there is as yet no published opinion on point. The defendant libraries have indicated that they believe the HathiTrust orphan works program is justified under fair use. *See* Defendants’ Joint Answer and Defenses at 23, *Authors Guild, Inc. v. HathiTrust*, No. 11 Civ. 6351 (S.D.N.Y. Dec. 2, 2011). Should this case result in a published opinion on the fair use question as it relates to orphans, it will likely shed substantial light on the issue of whether and how fair use might apply to making orphan works available within a research library community. As explained *supra* note †, the author is a faculty member at the

using entire works—are sometimes thought to be disfavored under the fair use doctrine, close scrutiny of the most comparable cases and of the purposes animating the fair use doctrine supports finding fair use for a variety of typical uses of orphan works by libraries and archives.

Part II of this Article briefly introduces the fair use concept and discusses its overall place in the debate over orphan works. Part III explains why the orphan status of a work is highly relevant to the fair use analysis. While the second factor of the fair use statute—the “nature of the copyrighted work”—often has little significance, Section III.A argues that this second factor should have a prominent role in orphan works cases. This prominent role is necessary because the work’s “nature” as an orphan both gives insight into the author’s likely incentives for creating the work and reveals how its orphan status fundamentally skews the market for the work. As such, orphan works are analogous to the unavailable “out of print” works that Congress and a handful of courts have described as favorable candidates for fair use under the second factor. Section III.B considers the fourth factor; the effect of the use on the market for the copyrighted work. This Section argues that the market factor should weigh decidedly in favor of fair use in the case of orphan works, due to a lack of cognizable harm to orphan owners and to the societal cost of preventing use of these works. Libraries’ and archives’ use of orphans would generate significant social value from which society should benefit. Orphan works, however, represent a clear market failure: there is no realistic possibility of completing a rights clearance transaction, no matter how high the costs of that transaction, because one party to the transaction is missing. This scenario creates a public loss, and militates in favor of fair use. For nonprofit library and archive uses in particular, it also militates against creating artificial markets via collective licensing, due to the socially wasteful misallocation of resources such efforts would require.

Part IV explains that the purposes animating library and archive uses of orphans will generally favor fair use because these uses further the constitutional goals of copyright. Section IV.A considers the important first fair use factor—the “purpose and character” of the use—and the favored uses described in the preamble to § 107. This Part explains that many purposes pursued by libraries and archives in digitizing and making available orphans are favored under both the preamble and the purpose factor, and go far to fulfill copyright’s constitutional purpose to promote progress more generally. Section IV.B contemplates the third fair use factor—the “amount

University of California, Berkeley, but does not represent, and does not speak, for any party involved in this case.

and substantiality” of the portion of the copyrighted work that was used. While the library and archive uses considered here would surely involve some entire works, a fact that is often thought to disfavor fair use, courts now assess the amount factor in light of the purpose for which the work is used. The socially beneficial purposes—such as preservation, research, and education—for which libraries and archives are likely to copy orphans will often demand reproduction of entire works. Combined with the fact that using entire orphan works is no more likely to harm a nonexistent market than use of only parts of those works, this factor should often be found, at worst, neutral to a defendant claiming fair use.

Part V describes additional considerations that may be helpful for libraries and archives that choose to rely on fair use. Part VI concludes.

Note that this Article takes no position about whether fair use is an appropriate solution for uses of orphan works outside of the nonprofit library and archival uses described above. For example, it does not consider making derivative works from orphan works, nor does it consider commercial uses at all. For some such uses, fair use may also appropriately play a role. For others, different proposed solutions—such as the original proposal by the U.S. Copyright Office to limit remedies for uses made after a reasonable search for the copyright owner—may instead be most appropriate.

II. BRIEF BACKGROUND ON ORPHAN WORKS AND FAIR USE

“Orphan works” are those presumptively copyrighted works for which the owner cannot be found.²⁴ There is broad agreement that orphans present

24. This definition is taken from the U.S. Copyright Office’s 2006 Report on Orphan Works. U.S. COPYRIGHT OFFICE, *supra* note 8, at 1. This definition is widely accepted by policymakers and academics. See, e.g., *Proposal for a Directive of the European Parliament and of the Council on Certain Permitted Uses of Orphan Works*, COM (2011) 289 (May 24, 2011), available at <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0289:FIN:EN:PDF> [hereinafter EU ORPHAN WORKS DIRECTIVE]; COMITÉ DES SAGES, *THE NEW RENAISSANCE* 16 (2011), available at http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/final_report_cds.pdf; EUROPEAN DIGITAL LIBRARIES, *FINAL REPORT ON DIGITAL PRESERVATION, ORPHAN WORKS, AND OUT-OF-PRINT WORKS* (2008), available at http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-clean171.pdf; IAN HARGREAVES, *DIGITAL OPPORTUNITY: A REVIEW OF INTELLECTUAL PROPERTY AND GROWTH* 38–40 (2011), available at <http://www.ipo.gov.uk/ipreview-finalreport.pdf>; Olive Huang, *U.S. Copyright Office Orphan Works Inquiry: Finding Homes for the Orphans*, 21 BERKELEY TECH. L.J. 265, 265 (2006); Bernard Lang, *Orphan Works and the Google Book Search Settlement: An International Perspective*, 55 N.Y.L. SCH. L. REV. 111, 116 (2010); Pamela Samuelson, *The Google*

a serious and growing issue for library and archive digitization projects.²⁵ In theory, any work could be an orphan, but the likelihood of orphan status increases with age. As time passes, information about the ownership of some works fades, and the task of connecting a work to its owner becomes more difficult, more costly, and in some cases impossible—making “clearing” the works by obtaining permissions also impossible. As copyright has changed over time, the likelihood of works becoming separated from owners has grown. For example, longer and longer copyright terms increase the likelihood that works will become disconnected from their owners. And as discussed below, in passing the 1976 Copyright Act, Congress recognized that the Act’s abolishment of copyright renewal requirements might increase the number of works covered by copyright even where there is lack of a commercial interest that would encourage their exploitation. Today, though it is difficult to determine the number of orphan works, it is clear that there are a great many, at least hundreds of thousands and perhaps many millions.²⁶

If these works cannot be digitized for preservation and for access by the public, then their “orphaning” creates a large social cost and a significant drag on the copyright system’s purpose to encourage the spread of knowledge. Many could disappear before they can enter the public domain and freely be repurposed. Liberating²⁷ these works from copyright limbo, on the other hand, would have great social benefit, both for today’s public and for tomorrow’s generations of scholars, students, and citizens. Accordingly, an international consensus has recently developed that there is an urgent need to address the orphan works problem.²⁸

Books Search Settlement as Copyright Reform, WIS. L. REV. 479, 483 (2011). *But see* Kahle v. Gonzales, 487 F.3d 697, 698 (9th Cir. 2007) (using “orphan” as plaintiffs suggested, to mean “works that allegedly have little or no commercial value but remain under copyright protection”).

25. See David R. Hansen, *Orphan Works: Causes of the Problem* 1 (Berkeley Digital Library Copyright Project, White Paper Nov. 3, 2012), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2038068; HARGREAVES, *supra* note 24; COMITÉ DES SAGES, *supra* note 24; U.S. COPYRIGHT OFFICE, *supra* note 8.

26. For example, estimates of the number of orphans in the Google Book Search corpus vary from a minimum of hundreds of thousands to several millions. See John P. Wilkin, *Bibliographic Indeterminacy and the Scale of Problems and Opportunities of “Rights” in Digital Collection Building* (Council on Library & Res. Ruminations No. 1, 2011), available at <http://www.clir.org/pubs/ruminations/01wilkin/wilkin.html/wilkin.pdf>.

27. See Lydia Pallas Loren, *Abandoning the Orphans: An Open Access Approach to Hostage Works*, 27 BERKELEY TECH. L.J. 1431 (2012).

28. See EU ORPHAN WORKS DIRECTIVE, *supra* note 24; HARGREAVES, *supra* note 24; COMITÉ DES SAGES, *supra* note 24.

There is less agreement on a proposed solution to the problem.²⁹ Ideas vary, but cluster around two main themes: legislative fixes that limit the remedies available to copyright owners who resurface even after a follow-on user has completed a “diligent search” and failed to find them; or license schemes administered by collecting societies or other central bodies.³⁰ These are generally all worthy ideas, though each has its limitations.³¹

This Article contends that the existing doctrine of fair use provides a way to broaden access to orphans. Fair use is one of the most distinctive features of U.S. copyright law.³² It stems from and supports copyright’s fundamental purpose to “promote the Progress of Science”³³ by balancing incentives to create original works with the social benefits that flow from the broad dissemination of those works, and by preserving fundamental First Amendment rights and free expression.³⁴

29. Proposed solutions to the orphan works problem are quite varied. Some, such as the Copyright Office’s remedies limitation approach, are quite simple. *See* U.S. COPYRIGHT OFFICE, *supra* note 8, at 115–21. Others, such as proposals for extended collective licensing or other transaction regimes, could prove to be highly complex, but they create additional certainty for users. *See generally* David R. Hansen, *Orphan Works: Mapping the Possible Solution Spaces* (Berkeley Digital Library Copyright Project, White Paper No. 2, 2012), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2019121 (describing a variety of proposals).

30. Hansen, *supra* note 29, at 1–2.

31. A full discussion of other proposed solutions is outside the scope of this Article. I discuss some of the limitations of licensing schemes *infra* Section III.A.3. Remedies limitations and registries or other formalities each have strong potential to effectively attack some aspects of the problem. Remedies limitations could effectively reduce the fear of downside risk, allowing more use of orphans *See, e.g.*, U.S. COPYRIGHT OFFICE, *supra* note 8, at 115–21. Formalities-type solutions could directly address the problem by preventing the creation of orphans in the first place: a system that encourages or requires registering a work and its owner in a central database could go far toward keeping works connected to their owners. A main limitation of each of these proposals is the political will that would be required to muster the necessary legislative changes; however, there are some promising signs in recent national and international discussions. *See* STEF VAN GOMPEL, *FORMALITIES IN COPYRIGHT LAW: AN ANALYSIS OF THEIR HISTORY, RATIONALES AND POSSIBLE FUTURE* (2011). In addition, the U.S. Copyright Office recently issued a new Notice of Inquiry on the orphan works issue.” *See supra* note 9.

32. Fair use has been a feature of U.S. copyright law since at least 1841, when Justice Story drafted *Folsom v. Marsh*, 9 F. Cas. 342 (C.C.D. Mass. 1841) (No. 4901), but the fundamental ideas behind it were introduced into copyright law by English courts of equity at least a hundred years earlier. *See, e.g.*, *Gyles v. Wilcox*, (1740) 26 Eng. Rep. 489 (Ch.) 490 (discussing “fair abridgment”). In the 1976 Copyright Act, Congress restated the doctrine in statutory form in § 107. 17 U.S.C. § 107 (2010).

33. U.S. CONST., art. I, § 8, cl. 8.

34. *Eldred v. Ashcroft*, 537 U.S. 186, 219–20 (2003) (characterizing fair use as providing “traditional First Amendment safeguards”). Fair use’s purpose in promoting and protecting expression is described in numerous cases. *See, e.g., id.*; *Campbell v. Acuff-Rose*

Unlike other exceptions and limitations in the Copyright Act,³⁵ and in most foreign copyright regimes,³⁶ fair use is not limited to specific actors or specific uses. Rather, it is an “equitable rule of reason,”³⁷ developed by courts and embodied in § 107’s familiar four factors: What is the purpose of the use? What is the nature of the allegedly infringed work? What is the amount and substantiality of what has been taken from the allegedly infringed work? And what is the effect of the use on the market for the allegedly infringed work?³⁸ Courts apply these factors on a case-by-case basis. Because of its equitable nature, fair use can flexibly accommodate changes in both copyrighted works and uses of those works over time.³⁹ Accordingly, fair use

Music, Inc., 510 U.S. 569 (1994); Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539 (1985); Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146, 1163 (9th Cir. 2007).

35. For example, § 108’s protections for certain library and archival uses are contingent on various requirements, including whether the library or archive is open to the public and how the library distributes the copies it makes, and § 110’s exception is limited to “face-to-face teaching” at a “nonprofit educational institution, in a classroom or similar place devoted to instruction.” 17 U.S.C. §§ 108, 110 (2010).

36. Other countries do have limitations protecting socially beneficial uses, but they tend to enumerate specific protected activities rather than using a generalized standard. *See, e.g.*, Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537, 2540 n.17 (2009) (citing the Council Directive 2001/29/EC, art. 5, On the Harmonization of Certain Aspects of Copyright and Related Rights in the Information Society, 2001 O.J. (L 167) 10 (establishing allowable exceptions and limitations to copyright in the EU)). For examples of more recent flexible limitations, see *infra* note 219.

37. H.R. REP. NO. 94-1476, at 65 (1976).

38. 17 U.S.C. § 107 (2010). In its entirety, § 107 reads:

Notwithstanding the provisions of §§ 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—

(1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;

(2) the nature of the copyrighted work;

(3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and

(4) the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.

Id.

39. *See, e.g.*, Sony Corp. v. Universal City Studios, Inc., 464 U.S. 417, 447–56 (1984) (accommodating “time-shifting” of television shows via video tape recorders); Sega Enters.

is especially well-suited to copyright problems that arise from evolving situations, such as the orphan works phenomenon.⁴⁰

It is thus unsurprising that both the Copyright Office and Congress took care not to disturb fair use when designing the 2008 proposed legislation for orphan works. The Copyright Office's report "stress[ed] that the presence of an orphan works provision should not act as a substitute or replacement for fair use."⁴¹ The bills introduced in both the House and Senate included savings clauses making clear that fair use would not be affected by orphan works legislation, and explicitly referred to the user of orphan works as the "infringer," rather than the "user" or some other term.⁴² This reference was, in part, an attempt to make clear that those engaging in fair uses of orphan works—who are not infringers—were not limited by the strictures of the proposed statute.⁴³ It also sought to clarify that the fair use doctrine overall would not be inadvertently limited by it.⁴⁴

v. *Accolade, Inc.*, 977 F.2d 1510, 1520–28 (9th Cir. 1992) (accommodating "reverse engineering" of computer programs); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 817–22 (9th Cir. 2003) (accommodating an image search engine).

40. As works in the collections of libraries and archives age, it seems likely that they will simultaneously become more likely to become "orphaned" and in greater need of preservation. At the same time, new methods of digitizing, organizing, and making works available are likely to arise, as are new uses of digital corpuses. Fair use's flexibility in accommodating such developments makes it, where it applies, especially appropriate for addressing orphan works issues over time.

41. U.S. COPYRIGHT OFFICE, *supra* note 8, at 56.

42. Shawn Bentley Orphan Works Act of 2008, S. 2913, 110th Cong. § 514(d) (2008); Orphan Works Act of 2008, H.R. 5889, 110th Cong. § 514(e) (2008); Orphan Works Act of 2006, H.R. 5439, 109th Cong. § 514(c) (2006).

43. For example, both the House and Senate 2008 bills required a user of an orphan work to designate the orphan work with a symbol, to provide attribution as possible, and to negotiate in good faith with later-appearing owners. S. 2913 § 514(b)(1); H.R. 5889 § 514(b)(1).

44. This observation proceeds directly from my participation in discussions around the proposed legislation; it is one interpretation that was offered during discussions among stakeholders. As an interpretation, this follows directly from the language of the Copyright Act. 17 U.S.C. § 107 (2006) (stating "[fair use] is not an infringement of copyright"); 17 U.S.C. § 106 (2006) (stating that the copyright owner's exclusive rights under Title 17 are "subject to" § 107). It is, however, important to note that the question of whether a fair user is an "infringer" is a matter of robust debate. PETER JASZI & PATRICIA AUFDERHEIDE, RECLAIMING FAIR USE: HOW TO PUT BALANCE BACK IN COPYRIGHT (2011). Further, as a matter of procedure, courts often treat fair use as an affirmative defense to an infringement—in other words, an excused infringement. *See Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 (1994); *Perfect 10 Inc. v. Amazon.com Inc.*, 508 F.3d 1146, 1163 (9th Cir. 2007); U.S. COPYRIGHT OFFICE, *supra* note 8, at 56 ("We stress that the presence of an orphan works provision should not act as a substitute or replacement for fair use.").

Indeed, in its discussion of the 1976 Copyright Act,⁴⁵ Congress both recognized the problems inherent in “t[ying] up a substantial body of material that is probably of no commercial interest” that would likely arise under the Act’s shift away from renewal requirements, and offered fair use as an ameliorator, stating in the House report:

It is true that today’s ephemera represent tomorrow’s social history, and that works of scholarly value, which are now falling into the public domain after 28 years, would be protected much longer under the bill . . . it is important to realize that the bill would not restrain scholars from using any work as source material or from making “fair use” of it⁴⁶

This dual recognition of a specific problem and the appropriateness of fair use in addressing it should further lead us to take fair use especially seriously in orphan works cases.

It also brings us to the broader question addressed in this Article: whether the orphan status of a work is especially relevant for fair use. The following discussion argues that such status is relevant, focusing first on the “nature” of an orphan work in a library or archive collection, and second on the follow-on effects of that orphan status for fair use’s market-harm analysis.

III. THE ORPHAN STATUS OF WORKS IS HIGHLY RELEVANT TO THE NATURE-OF-WORK AND MARKET-HARM FACTORS AND FAVORS FAIR USE

A. THE “NATURE” OF AN ORPHAN WORK

The U.S. copyright law’s fair use provision directs courts to consider “the nature of the copyrighted work” that the plaintiff claims was infringed.⁴⁷ According to Judge Leval, this factor “concerns itself with protecting the incentives of authorship. It implies that certain types of copyrighted material are more amenable to fair use than others.”⁴⁸ In evaluating the nature of a work, courts have sometimes limited the inquiry to whether the work is more

45. Copyright Act of 1976, Pub. L. No. 94-553, 90 Stat. 2541 (codified as amended at 17 U.S.C. §§ 101–805).

46. H.R. REP. NO. 94-1476, at 136 (1976). As the Copyright Office noted in its *Report on Orphan Works*, “Congress recognized the problem [with extending copyright terms and abolishing renewal requirements], but considered it to be outweighed by the many benefits of the new system,” such as removing the “traps for the unwary” created by formalities requirements. U.S. COPYRIGHT OFFICE, *supra* note 8, at 43.

47. 17 U.S.C. § 107(2) (2010).

48. Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1116 (1990).

factual or more creative, and whether it has been published or remains unpublished.⁴⁹ The usual outcome of this inquiry is that fair use of a work is easier to establish when the work is factual rather than creative,⁵⁰ and when it has been published rather than remaining unpublished, though the latter depends heavily on the factual context.⁵¹ In addition, courts have sometimes focused the nature factor on whether the work was out-of-print and unavailable on the market.⁵²

It is unusual to begin a fair use analysis with this factor. The nature-of-work factor is widely understood to weigh lightly in most fair use cases, and courts often truncate their discussion of this factor⁵³ so that it has little significance in the context of the broader analysis.⁵⁴ But the nature factor is important in some cases, notably *Harper & Row v. Nation Enterprises*, in which the work's unpublished status was tied to a valuable interest in first serialization,⁵⁵ and in *Sega v. Accolade* and related cases, in which the software programs at issue contained unprotected ideas that “afford[ed] them a lower degree of protection than more traditional literary works.” Additionally, some commentators, notably Pierre Leval and Robert Kasunic, have proposed that the nature-of-work factor should receive deeper consideration more generally.⁵⁶ Judge Leval argues that inquiry into the nature of the copyrighted work should give insight into copyright's goal “to stimulate authorship for the public edification” by “determ[in]ing whether the work is the type of material that copyright was designed to stimulate, and whether the secondary

49. See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 586 (1994); *Stewart v. Abend*, 495 U.S. 207, 237–38 (1990); *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 563–64 (1985); Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions: 1978–2005*, 156 U. PA. L. REV. 549, 610–15 (2008).

50. Beebe, *supra* note 49, at 611.

51. *Harper & Row*, 471 U.S. at 564; Beebe, *supra* note 49, at 612–15.

52. *Harper & Row*, 471 U.S. at 553 (“If the work is ‘out of print’ and unavailable for purchase through normal channels, the user may have more justification for reproducing it.”) (quoting S. Rep. No. 94-473, at 64 (1975)); cf. *Peter Letterese & Assocs., Inc. v. World Inst. of Scientology Enters.*, 533 F.3d 1287, 1313–14 (11th Cir. 2008). *But see* *Basic Books, Inc. v. Kinko's Graphics Corp.*, 758 F. Supp. 1522, 1533 (S.D.N.Y. 1991).

53. See Robert Kasunic, *Is That All There Is? Reflections on the Nature of the Second Fair Use Factor*, 31 COLUM. J.L. & ARTS 529, 529 (2008); Leval, *supra* note 48, at 1116 (“The nature of the copyrighted work is a factor that has been only superficially discussed and little understood.”). According to Barton Beebe, “[d]espite § 107's command that ‘the factors to be considered shall include’ factor two, 17.7% of the . . . opinions [applying § 107] failed even to refer to the factor, while an additional 6.5% did so only to call it irrelevant.” Beebe, *supra* note 49, at 610.

54. MELVILLE B. NIMMER & DAVID NIMMER, *NIMMER ON COPYRIGHT* § 13.05 (2011).

55. *Harper & Row*, 471 U.S. at 564 (noting the effect on plaintiff's right of first serialization meant that the unpublished nature of the work weighed heavily against fair use).

56. Leval, *supra* note 48, at 1116–22; Kasunic, *supra* note 53, at 530.

use . . . would interfere significantly with the original author's entitlements."⁵⁷ Kasunic agrees, though he argues that a better form of this question is "whether copyright might have reasonably encouraged or provided an incentive for an author to create the work" taking into account the norms and customs of the relevant genre.⁵⁸

This more probing approach to the nature-of-work factor has much to recommend it, especially in orphan works cases. In general, such an inquiry can illuminate whether copyright's creation and dissemination goals are furthered by allowing the copyright owner full control or instead allowing fair use, and relatedly, whether the defendant's use is likely to invade a "traditional, reasonable, or likely to be developed" market under the market-harm factor.⁵⁹ In the specific case of orphan works, the fact that a work has been orphaned—that is, disconnected from its owner and thus made unavailable on the market—could easily be considered as one important aspect of the work's "nature" in fair use cases. First, the "orphan" aspect of a work's nature strongly suggests that it was not sufficiently available on the market to support copyright's dissemination goals. As such, considering orphan status under this factor would be analogous to some courts' treatment of works whose nature includes being unavailable on the market.⁶⁰ Second, recognizing the "orphan" aspect of a work in question should prompt additional inquiries into the work's nature, such as whether or not it was created in response to copyright incentives. While the fact that a work is an orphan is not by itself sufficient to conclude whether a work was created in response to copyright incentives, such status is evidence that should provoke further inquiry. Some relevant inquiries are described in more detail below.

Beginning with the nature-of-work factor here can thus give insight into the importance for fair use cases of a work's orphan status. More broadly, this factor can also shed light on the importance of a variety of other characteristics of the "nature" of orphan works that are likely to be in library and archival collections. Overall, a work's "orphaned" status provides a distinct and strong argument for giving considerable weight to this factor in any case involving orphans. The strength of this argument is intensified by other aspects of the nature of many works in library and archive collections. Armed with insight from this inquiry, it is then straightforward to proceed to considering the market-harm and purpose questions.

57. Leval, *supra* note 48, at 1118–19.

58. Kasunic, *supra* note 53, at 540.

59. *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930 (2d Cir. 1994).

60. *See infra* note 69 and accompanying text.

1. *Orphan Status as a Key Aspect of a Work's "Nature"*

The signature characteristic of an orphan work's "nature" is that its owner cannot be found.⁶¹ Whether the missing owners are defunct organizations, anonymous apparatchiks, or simply individual creators lost in the sands of time, they are by definition no longer tied to the work,⁶² and it is

61. A full discussion of how best to identify orphans is outside of the scope of this Article. In brief, however, there is broad consensus around conducting a threshold "diligent search" for copyright owners. COPYRIGHT OFFICE, *supra* note 8, at 5–6 ("Almost every commenter who advocated a limitation-on-remedies system agreed that a fundamental requirement for designation of a work as orphaned is that the prospective user have conducted a search for the owner of the work, and that the search results in the owner not being located. The commenters differed in the types of searches they would consider adequate."); Shawn Bentley Orphan Works Act of 2008, S. 2913, 110th Cong., at 6–7 (2008); EU ORPHAN WORKS DIRECTIVE, *supra* note 24, at 10. Details about what might constitute a reasonably diligent search are a topic for another paper. For the purposes of this Article, I have assumed that "orphan" means that a search that is reasonable under the circumstances has occurred, and no owner has been found.

However, it is important to note that, without care, search costs could easily become unreasonably high, quashing the incentive or ability to make a socially beneficial use. Search costs could, in fact, quickly become exceptionally burdensome, as a search for an unlocatable owner is, at a basic level, an attempt to prove a negative (that no one exists or that the identified owner cannot be found) that may be unprovable. Any solution to the orphan works problem, including fair use, cannot require searches so onerous as to undermine the benefits of the solution by making costs so high as to dampen incentives to use orphans or prohibit the possibility entirely. This is obviously especially important for nonprofit libraries and archives, which have limited budgets, forcing them to be cost-sensitive.

62. Of course, if a court is considering fair use, then this means that a putative owner has appeared and claimed the work; as such, it may no longer be an orphan. This occurrence prompts two observations. First, courts should take care to address the question of ownership early in the proceedings. In infringement cases, copyright holders bear the burden of proving that they own the work in question, and it will be particularly important to address this question for an owner who was not locatable after a reasonable search. Plaintiffs do make mistakes in claiming ownership. *See, e.g., Cambridge Univ. Press v. Becker*, 863 F. Supp. 2d 1190 (N.D. Ga. 2012) (holding that plaintiffs could not prove ownership for the works at issue in multiple claimed infringements). Second, assuming that the plaintiff is indeed the correct owner, then the effect of the work's orphan status on the application of the fair use doctrine should be considered for any use of the work up until the owner reappeared to claim the work. Once the owner is back in the picture, then the question becomes whether the use should be licensed or still is fair. Note also that plaintiffs have already attempted to address orphan works even where the owner has not come forth. While they were unsuccessful in doing so, the *Authors Guild v. HathiTrust* organizational plaintiffs claimed associational standing on behalf of all of their members. This scenario is another way orphan owners might "appear" before a court. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *4–7 (S.D.N.Y. Oct. 10, 2012). Plaintiffs may also attempt to bring class actions on behalf of a class that includes putative orphan owners. *See, e.g., Authors Guild, Inc. v. Google, Inc.*, Case No. 05 CV (S.D.N.Y. Feb. 13, 2012).

likely⁶³ that the work is no longer—if it ever was—tightly connected to their economic interests. Courts should recognize the nature-of-work factor’s value in informing their understanding of at least two key attributes of orphan works: first, that their participation in a market that might be harmed by allowing fair use is unlikely, at best; and second, whether the likely motivations behind the work’s creation stemmed from copyright’s economic incentives.

During the Copyright Office proceedings on orphan works, commenters pointed out that the fact that a work’s owner is missing will often indicate that the work has low, or no, market value to those owners.⁶⁴ After all, a work that cannot be connected to its owner after a diligent search has a high probability of having been economically abandoned.⁶⁵ This feature of orphan works is relevant to whether the use would “interfere significantly with the original author’s entitlements”⁶⁶ and whether such use is likely to invade a “traditional, reasonable or likely to be developed”⁶⁷ market under the market-harm factor. As discussed in more detail in Section III.B, *infra*, where a work

63. There are no doubt exceptions. Photographs, for instance, routinely lack author information. During the discussions about U.S. orphan works legislation, professional photographers expressed concern that orphan works legislation could unjustly deprive them of their ability to appropriately control the copyrights in their photographs because they are easily disassociated from the works. Response of American Society of Media Photographers to Orphan Works Notice of Inquiry, Mar. 25, 2005, Comment OW0668 at 1, *available at* <http://www.copyright.gov/orphan/comments/OW0668-ASMP.pdf>; Response of Professional Photographers of America to Orphan Works Notice of Inquiry, Mar. 25, 2005, Comment OW0642 at 1, *available at* <http://www.copyright.gov/orphan/comments/OW0642-PPA.pdf>. As better search mechanisms—especially visual matching technologies—develop, this understandable concern may be addressed over time. More generally, in such cases courts should conduct an inquiry into the characteristics of the work that could illuminate the reason for its lack of connection with an owner. While professional photographs may have economic value, untold thousands of amateur “snaps” do not, and were very likely created in response to motivations other than copyright incentives. In some cases, seeing the difference will be straightforward—for example, an industrial photograph created by a defunct company, apparently orphaned because the company no longer exists and most likely never had a copyright-based interest in the photo—some cases less so. The equitable inquiry of fair use is especially well-suited to such tasks.

64. *See generally* Response of Public Knowledge to Orphan Works Notice of Inquiry, Mar. 25, 2005, Comment OW0629 at 8, *available at* <http://www.copyright.gov/orphan/comments/OW0629-PublicKnowledge.pdf>; COPYRIGHT OFFICE, *supra* note 8, at 84.

65. Loren, *supra* note 27; Aryeh L. Pomerantz, *Obtaining Copyright Licenses by Prescriptive Easement: A Solution to the Orphan Works Problem*, 50 JURIMETRICS J. 195 (2010); Matthew W. Turetzky, Note, *Applying Copyright Abandonment in the Digital Age*, 2010 DUKE L. & TECH. REV. 19 (2010); Megan L. Bibb, Note, *Applying Old Theories to New Problems: How Adverse Possession Can Help Solve the Orphan Works Crisis*, 12 VAND. J. ENT. & TECH. L. 149 (2009).

66. Leval, *supra* note 48, at 1118–19.

67. *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930 (2d Cir. 1994).

has been left orphaned, such an invasion is very unlikely. Relatedly, orphan status indicates that the work is unlikely now to be disseminated to the public by its owner. As such, if libraries and archives can make the works available through fair use, it will advance knowledge without undermining markets for the works.

This analysis is in line with some courts' treatment of works that are unavailable on the market, as orphans by definition are. In the Senate Report accompanying the 1976 Act, Congress indicated that where a work "is 'out of print' and unavailable for purchase through normal channels, the user may have more justification for reproducing it than in the ordinary case."⁶⁸ A handful of courts have followed this reasoning in applying the nature-of-work factor to works unavailable on the market.⁶⁹ The same reasoning applies to orphans, and perhaps with additional force to those orphans that were never presented to the public in the first place.⁷⁰ The caselaw does rightly distinguish between works that no longer enjoy economic value or

68. S. Rep. No. 94-473, at 64 (1975).

69. *See, e.g.*, *Hofheinz v. A & E Television Networks*, 146 F. Supp. 2d 442, 447–48 (S.D.N.Y. 2001) (applying this reasoning to the use of a trailer for a film released almost fifty years earlier and not itself available on the market and finding that the second factor favored fair use); *Maxtone-Graham v. Burtchaell*, 803 F.2d 1253, 1264 n.8 (2d Cir. 1986) (stating that "[a] key, though not necessarily determinative, factor in fair use is whether or not the work is available to the potential user" and citing the Senate Report); S. Rep. No. 94-473, at 64 (1975) ("If the work is 'out of print' and unavailable for purchase through normal channels, the user may have more justification for reproducing it than in the ordinary case."); H.R. Rep. No. 94-1476, at 67 (1976). *But see Am. Geophysical Union*, 60 F.3d at 931.

70. Arguably, the public's loss is greater when works are never available at all than when some copies—even if they are rare—have been injected into commerce. Of course, the countervailing argument, which held sway in courts until *Campbell v. Acuff-Rose* and Congress' amendment of § 107, is that follow-on users' fair use rights are circumscribed in unpublished works, due to copyright holder's interest in determining when first publication should occur. Judge Leval and others roundly criticized over-weighting a work's unpublished status in the fair use analysis and, following *Campbell*, courts began to engage in more careful analysis of this attribute of a work. Leval, *supra* note 48, at 1122 ("The fact that a document is unpublished should be of small relevance unless it was created for or is on its way to publication."); *cf.* 17 U.S.C. § 107 (2010) ("The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the [fair use] factors."); *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 577–78 (1994) (stating that "the task is not to be simplified with bright-line rules," and "[a]ll [factors] are to be explored, and the results weighed together, in light of the purposes of copyright"). *But see Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 564 (1985) (noting that "the author's right to control the first public appearance of his expression weighs against such [fair] use of the work before its release" and stating that "[a] use that so clearly infringes the copyright holder's interests in confidentiality and creative control is difficult to characterize as 'fair'"). For orphans, specifically, as no owner can be found to give permission to publish, the public's loss appears particularly great, with little to no corresponding harm to the owner.

whose authors have left them to languish⁷¹ and works that have been purposely withheld from the market by the copyright owner.⁷² This latter set of works enjoys the full protection of copyright, and their unavailability on the market will not weigh in favor of fair use.⁷³ This latter reasoning might apply, for example, where a work's owner is locatable but recalcitrant or unresponsive. In such cases, however, the work is not orphaned.

Along with other aspects of its "nature," the "orphaned" status of a work may also provide insight into whether traditional copyright incentives supplied an impetus to create the work, and thus whether allowing fair use would damage future incentives. Of course, exactly how much a work's orphan status reveals about the creation impetus will depend on the entire set of specific characteristics of that work. However, given the fact that copyrights in them have been left to languish, large subsets of orphan works—such as those described in the examples given below—are likely to have been created for non-economic reasons.⁷⁴ As such, a work's orphan status should at least trigger an inquiry into whether its orphaning is connected to the motivations behind its creation.

This is an important question to consider because many incentives to create are unlikely to be injured by allowing fair use. Nor, as Judge Leval points out, is copyright law designed to provide every incentive that might lead an author to create. Leval argues that the nature-of-work factor should favor the original creator more for works created for reasons "at the heart of the purpose of copyright"—such as those works prepared for publication—than "in the case of a document written for reasons having nothing to do with the objectives of copyright law."⁷⁵ Kasunic further points out that inquiries into motivation and the author's "reasonable and customary expectations"⁷⁶ for a work's scope of protection can illuminate the appropriate leeway for fair use. For example, more factual works are reasonably expected to enjoy "thinner" copyright protection than creative

71. See, e.g., *Hofbeinz*, 146 F. Supp. 2d at 447–48.

72. See, e.g., *Peter Letterese & Assocs., Inc. v. World Inst. of Scientology Enters.*, 533 F.3d 1287, 1313–14 (11th Cir. 2008).

73. See *id.*

74. For some orphan works, of course, orphan status will be less informative on this point. Some orphans were created originally as commercial works, and have simply lost commercial value or had other misfortunes befall them over time. Courts should therefore consider the "orphan" aspect of a work's nature along with other characteristics in order to best analyze the work's orphan status under the nature factor.

75. Leval, *supra* note 48, at 1117.

76. Kasunic, *supra* note 53, at 541.

works; thus, a creator's copyright-based motivations are likely to be less affected by unpermitted use of a factual work than of a creative work.⁷⁷

To illustrate, Judge Leval uses the memorable example of an extortion note demanding “[p]ay me by Friday or I’ll break your goddamn arms”⁷⁸ as a work that would require no copyright incentives to stimulate creation or, indeed, any encouragement that copyright ought to provide.⁷⁹ As such, the extortionist neither needs copyright-based incentives nor has cognizable “reasonable expectations” of ownership rights to protect,⁸⁰ and the scope of fair use should be correspondingly broad. While Judge Leval’s extortion note is a somewhat fanciful example, orphan works residing in library and archive collections provide abundant real-world examples of the same principles.

For example, research libraries are replete with academic works created and published, not primarily to profit from copyright, but to share knowledge, find truth, and advance the scholar’s reputation and career;⁸¹ and with special collections that contain business records, personal correspondence, and many other items that were not prepared with copyright protection in mind. The Prelinger Archives, Internet Archive, and subject-matter archives like the LOC American Memory projects and the Civil Rights Movement Veterans project collectively preserve and make accessible a wide variety of ephemera, including corporate training films,⁸² military propaganda,⁸³ home movies,⁸⁴ amateur snapshots,⁸⁵ advertising copy,⁸⁶

77. *Id.* at 541–42.

78. Leval, *supra* note 48, at 1116–17.

79. *Id.* at 1119.

80. *See* Bill Graham Archives v. Dorling Kindersley Ltd., 448 F.3d 605, 612 (2d Cir. 2006) (quoting Leval, *supra* note 48, at 1117).

81. For one discussion of some academic authors’ self-described incentives, see Academic Author Objections to Plaintiff’s Motion for Class Certification at 1, Authors Guild, Inc. v. Google, Inc., Case No. 05 CV (S.D.N.Y. Feb. 13, 2012):

We write scholarly works on a regular basis. Our primary motivation in preparing these works is to share the knowledge we have cultivated with other scholars and interested members of the public. Although we are not indifferent to revenue streams we receive from books that we publish, the main reward we wish to attain from our intellectual labors is the satisfaction of contributing to the ongoing dialogue about issues of concern to us and, perhaps as an added bonus, a reputation for excellence in scholarship among our peers. A number of us have made some or all of our academic work available on an open access basis through Creative Commons licenses and the like.

82. Eric J. Schwartz, Founding Director/Board Member, National Film Preservation Foundation, Panel at Berkeley Orphan Works Conference: Who Wants to Make Use of Orphan Works and Why?, at slide 6 (Apr. 12, 2012) (slides available at http://www.law.berkeley.edu/files/paul_schwartz.pdf).

83. *Id.*

meeting minutes,⁸⁷ letters,⁸⁸ brochures,⁸⁹ marketing materials,⁹⁰ and countless other items. These examples were presumably made, respectively, to improve industrial processes, gain support and funding for military endeavors, record family experiences, sell products and services, memorialize organizational decisions, correspond, and advocate and inform, among other reasons—but not, in general, to reap economic gains made possible by copyright protection. Today, while the private value of these works may still be low, the societal value of these works is high.

At the recent conference on orphan works held at UC Berkeley, librarians and archivists gave especially trenchant examples of orphan works with high societal value and a low likelihood of having been made with copyright incentives in mind. Rick Prelinger of Prelinger Archives showed footage of “release” day at a World War II internment camp for American citizens and residents of Japanese descent. The smiles and waves of the interns as they leave the camp in military trucks contrasts with the societal shame of the moment. The filmmaker’s motivation might have been bureaucratic record-keeping, or perhaps more general historical recordation; there is no way to know. Despite the film’s present historical importance, the owner is long lost. Bruce Hartford of the Civil Rights Movement Veterans website showed a photograph of James Forman, head of SNCC, gazing out from within a Southern jail cell. The photographer’s identity is unknown, but his motivations are fairly clear. Former SNCC workers recall that he was also a prisoner in the jail; SNCC workers were not allowed to photograph in the prison, so they passed him a camera and he took the picture in response to a request.⁹¹ In both cases, the fact that the works have been left orphans supports an analysis that the motivations which provided the impetus for creation bore a limited relation, if any, to copyright incentives.

Overall, although a number of orphan works in library and archive collections were likely created with commercial incentives in mind, another number—perhaps a substantial number—are likely to represent the result of

84. *Id.*

85. Hartford, *supra* note 4, at slide 6.

86. *Prosperity and Thrift: The Coolidge Era and the Consumer Economy, 1921–1929*, LIBRARY OF CONGRESS (Aug. 9, 1999), <http://memory.loc.gov/ammem/coolhtml/coolhome.html>.

87. Hartford, *supra* note 4, at slide 6.

88. *Id.* at slide 7.

89. Hartford, *supra* note 4, at slide 6.

90. Kenneth Crews, Director, Columbia Copyright Advisory Office, Panel at Berkeley Orphan Works Conference: Who Wants to Make Use of Orphan Works and Why?, at slide 7 (Apr. 12, 2012) (slides available at <http://www.law.berkeley.edu/11731.htm>).

91. Hartford, *supra* note 4, at slide 6.

a wide array of non-copyright-based motivations to create. While owners may stay connected to works for a variety of noneconomic reasons, such as reputational value or sentiment, it seems likely that noneconomic motivations to create, and the attendant lack of ongoing economic interest in the ensuing works, contributed to the orphaning of many of these works. As such, inquiry into characteristics of the works' natures—including the fact that they are orphaned—is likely to provide useful guidance to courts deciding whether or not the work was likely created in response to copyright incentives. While orphan status alone is certainly not sufficient to conclude that a work was created in response to non-copyright-based motivations, orphan status should at least suggest looking further.

2. *Other Characteristics of Works in Library and Archive Collections*

As described above, the orphan status of a work, considered in concert with other characteristics of the work's "nature," is likely to yield information that is useful to a fair use analysis. There are, however, other reasons for courts to consider the nature of these works carefully in any fair use analysis. Many orphan works housed in libraries and archives are likely to exhibit additional characteristics of "nature"—for example, being highly factual rather than creative—that would likely weigh in favor of fair use irrespective of orphan status, and in addition to any evidence that inquiring into the work's nature presents about the author's motivation to create. In his *HathiTrust* order, for example, Judge Baer cited the fact that, on the whole, the HathiTrust Digital Library contained only about nine percent "prose fiction, poetry, and drama."⁹²

While the relevant characteristics will obviously vary from work to work and from collection to collection, one might expect to find a wide range of characteristics that would be considered favorable to fair use in library and archive collections. The scholarly works mentioned above will often be more factual than creative in nature, and this is (perhaps surprisingly) true of most other works in libraries as well.⁹³ Special collections and subject-matter based archives contain large amounts of material, such as the ephemera described above, that have little economic value but a great deal of historical and cultural value. Many of these works, such as those in the subject-matter-

92. *Authors Guild, Inc. v. HathiTrust*, No. 11 Civ. 6351, slip op. at 18. Of course, this set is not a perfect proxy for "creative" work in the collection, but it is a suggestive indication.

93. A recent study of post-1923 works in WorldCat shows that fiction only makes up about eight percent of library holdings. See Brian Lavoie & Lorcan Dempsey, *Beyond 1923: Characteristics of Potentially In-Copyright Print Books In Library Digital Collections*, 15 D-LIB MAG. (Nov./Dec. 2009), available at <http://www.dlib.org/dlib/november09/lavoie/11lavoie.html>.

based American Memory collections, derive a significant portion of that value from the fact that they are contained within a curated collection; these works would be much less valuable on their own.⁹⁴ The rare, obscure, or ephemeral works contained in library and archive collections would likely *never* be available to the public if it were not for those organizations' missions to collect, preserve, and provide access to them.⁹⁵ All of these observations stem from considering the characteristics of these works in more detail than is generally done in fair use cases, lending further credence to Leval's and Kasunic's argument that the nature-of-work factor should be carefully considered.

In sum, serious inquiry into the "nature" of orphan works in library and archive collections is likely to reveal useful information about the appropriate breadth for fair use in any given dispute. Courts would be well-served by considering the nature-of-work factor carefully in the case of orphan works being digitized and made available by libraries and archives. In each example given above, consideration of the work's orphan status and other evidence of the motivations behind its creation and its creator's "reasonable and customary" expectations for its protection⁹⁶ could help courts analyze the equities under § 107. For orphans, the equities will often weigh in favor of allowing preservation, access, and dissemination under fair use, as copyright incentives are unlikely to be damaged by the use of works that were not created in response to those incentives in the first place. And as noted above, inquiry under the nature-of-work factor is likely to provide insight into the best analysis under § 107's fourth factor—"the effect of the use upon the potential market for or value of the copyrighted work,"⁹⁷ to which we turn next.

3. *The Market-Harm Factor and Orphan Works*

One of the most important factors that courts consider when analyzing fair use is the effect that the defendant's use will have on markets for the copyrighted work.⁹⁸ This factor's importance stems from its close connection

94. For example, one set of meeting minutes may hold little interest, but a series of meeting minutes that reveal an organization's development over time may have high historical value.

95. It is important to note here that many works in special collections are likely to be unpublished. *See supra* note 70 and accompanying text (discussing the weight given in the fair use analysis to a work's unpublished status).

96. Kasunic, *supra* note 53, at 541.

97. 17 U.S.C. § 107(4) (2010).

98. *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 566 (1985); Matthew Sag, *Predicting Fair Use*, 73 OHIO ST. L.J. 47, 63–64 (2012); Beebe, *supra* note 49, at 616–18. For some time, lower courts applied the Supreme Court's discussion of the market

to the incentive theory underlying U.S. copyright law. A use that does not harm the potential market for a work does not diminish the incentive to create that copyright provides, and is more likely to be fair use.⁹⁹ Where the defendant's use of a copyrighted work harms an actual or reasonably foreseeable potential market for the work, incentives may be damaged, and the market factor will likely weigh against fair use.¹⁰⁰

The fourth factor is an obviously salient factor in evaluating the fair use defenses as to orphan works. Orphan works are unusual in that they represent a truly complete market failure.¹⁰¹ Where there is no locatable owner with whom to transact—in other words, where one party to any transaction is simply missing¹⁰²—there cannot be a negative effect on the market for the work because no market can arise.¹⁰³ Accordingly, while the potential societal benefit of transactions may be great, they cannot realistically occur.

factor in *Harper & Row*, 471 U.S. at 566 (“This last factor is undoubtedly the single most important element of fair use.”). See, e.g., *Sega Enters. v. Accolade, Inc.*, 977 F.2d 1510, 1523 (9th Cir. 1992); *Princeton Univ. Press v. Michigan Document Services, Inc.*, 99 F.3d 1381, 1385 (6th Cir. 1996). Such an approach, however, could undermine the overall purpose of § 107's protections for free expression and other salutary values, and the Court clarified its approach in *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 577–78 (1994) (stating that “[t]he task is not to be simplified with bright-line rules,” and “[a]ll [factors] are to be explored, and the results weighed together, in light of the purposes of copyright”). Since then, the importance of the “purpose of the use” has grown and solidified. See *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 608–12 (2d Cir. 2006); see also *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818–20 (9th Cir. 2002).

99. *Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417, 450 (1984) (noting that “a use that has no demonstrable effect upon the potential market for, or the value of, the copyrighted work need not be prohibited in order to protect the author's incentive to create”).

100. The leading articulation of this reasoning is the Second Circuit's in *American Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 929–31 (2d Cir. 1994).

101. For another example of market failure in the copyright context, see Wendy Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and its Predecessors*, 82 COLUM. L. REV. 1600, 1621 (1982). Gordon writes:

When a new use for copyrighted works becomes available to the public, market mechanisms may take time to develop. At early stages of use, the transaction costs that would be involved for a user to purchase permission to use, or for the copyright owner to seek enforcement against nonpaying users, might well exceed whatever gain the parties might otherwise expect from the transaction. A custom therefore may develop under which users proceed without permission.

Id.

102. Identifying orphans may, of course, be challenging. Assuming, however, that a reasonable search has been made for the owner, market failure can reasonably be assumed.

103. For a discussion of the possibility of *creating* a market of sorts through a collective licensing scheme, see *infra* notes 126–144 and accompanying text.

Given this, harm to an owner's market from a follow-on use is vanishingly unlikely for true orphans. This circumstance should weigh heavily in favor of fair use where libraries and archives plan to digitize orphans. It also prompts two further observations. First, in order to create reasonable certainty that the market failure is as profound as asserted,¹⁰⁴ orphans must be properly identified.¹⁰⁵ Second, where orphans are properly identified and the purpose of the use is appropriate, a licensing scheme for orphan works is not only unnecessary under the fair use doctrine, but would represent a highly inefficient misallocation of resources.¹⁰⁶

B. MARKET HARM

In considering the likelihood of harm to the market for a copyrighted work, courts consider both what the relevant market or markets for the work might be, and whether the defendant's use of the work is likely to cause harm to those markets. Broadly speaking, courts weigh the market-harm factor in favor of fair use where no relevant market exists or where the defendant's use does not invade a relevant market. Not all theoretically possible markets are relevant to the analysis; in the leading case on this point, the Second Circuit explains that "only traditional, reasonable, or likely to be developed markets" should be considered.¹⁰⁷ In formulating this test, the Second Circuit followed the Supreme Court's descriptions of relevant markets as "only those that creators of original works would in general develop or license others to

104. For non-orphans, the transfer of rights needed to make a socially desirable use may naturally occur through negotiation and mutual agreement between the user and the copyright holder. See Gordon, *supra* note 101, at 1615. Note, however, that, while it may not be as complete if works are not true orphans, there may still be substantial market failure in the form of transaction costs that substantially outweigh the benefits of transacting, especially for older or obscure works, and especially for mass digitization projects. See, e.g., *id.* at 1628–30 (discussing transaction costs as a barrier where otherwise a market would function effectively); Samuelson, *supra* note 36, at 2614 (noting that "another important factor [weighing against copyright claims in *Field v. Google*, 412 F. Supp. 2d 1106 (D. Nev. 2006)] was the exceptionally high transaction costs that Google would have incurred if it had to seek and obtain permission for every copy its web crawlers made of Internet content.").

105. See U.S. COPYRIGHT OFFICE, *supra* note 8, at 8 (discussing the consensus view that orphans should be designated where an owner is either unidentifiable or unlocatable after a "diligent search").

106. See *infra* text accompanying notes 116–23; Ariel Katz, *The Orphans, the Market, and the Copyright Dogma: A Modest Solution for a Grand Problem*, 27 BERKELEY TECH. L.J. 1285 (2012); Randal C. Picker, *Private Libraries and Orphan Works*, 27 BERKELEY TECH. L.J. 1259 (2012).

107. *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930 (2d Cir. 1994).

develop”¹⁰⁸ and “the normal market for a copyrighted work.”¹⁰⁹ Courts have found a lack of “traditional” or “normal” markets where, for example, the use is critical¹¹⁰ or otherwise transformative,¹¹¹ or where a license to the work in an appropriate format is not offered for a reasonable price.¹¹² Where licenses are readily available for the defendant’s use, on the other hand, a viable relevant market is likely to exist.¹¹³

Courts are less likely to find market harm where a market fails to form or has socially harmful defects, such as inordinate transaction costs or negative externalities.¹¹⁴ Markets can fail for a wide range of reasons. A classic

108. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 592 (1994) (“The market for potential derivative uses includes only those that creators of original works would in general develop or license others to develop.”).

109. *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 568 (1985) (fourth factor concerned with “use that supplants any part of the normal market for a copyrighted work”) (quoting S. Rep. No. 94-473, at 65 (1975)).

110. *See, e.g., Campbell*, 510 U.S. at 592 (“[T]he law recognizes no derivative market for critical works.”).

111. *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 614–15 (2d Cir. 2006).

112. *Authors Guild, Inc. v. HathiTrust*, No. 11 Civ. 6351, slip op. at 20 (rejecting argument that a claimed potential market for mass digitization of library books showed harm under the market factor where the claimed potential market was ‘conjecture’); *Cambridge Univ. Press v. Becker*, 863 F. Supp. 2d 1190, 1237 (N.D. Ga. 2012) (“For loss of potential license revenue to cut against fair use, the evidence must show that licenses for excerpts of the works at issue are easily accessible, reasonably priced, and that they offer excerpts in a format which is reasonably convenient for users.”).

113. *See, e.g., Harper & Row*, 471 U.S. 539, 602–03 (weighing the fourth factor against fair use where the plaintiff was able to show a direct loss of licensing revenue for the infringed work and where there was substantial potential for damage to the market for first serialization rights even if direct loss had not been shown); *Becker*, 863 F. Supp. 2d at 1255 (weighing the fourth factor against fair use where a license in the correct digital format was readily available); *United States v. Am. Soc’y of Composers, Authors & Publishers*, 599 F. Supp. 2d 415, 432–33 (S.D.N.Y. 2009) (“The existence of these markets [for derivative works of original ASCAP music] militates against a finding of fair use.”).

114. *See, e.g., Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417, 451–55 (1984) (holding that plaintiffs did not show home time-shifting harmed the market); *Kelly v. Arriba Soft Corp.*, 336 F. 3d 811, 821–22 (9th Cir. 2002); *Becker*, 863 F. Supp. 2d at 1290 (“[A]bsence of a ready market shifts the factor four analysis to favor fair use.”); *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1122 (D. Nev. 2006) (holding there was no evidence a “market for licensing search engines to access Web pages through ‘cached’ links” existed or was likely to develop); *Time Inc. v. Bernard Geis Assocs.*, 293 F. Supp. 130, 146 (S.D.N.Y. 1968). Courts do not necessarily discuss such cases in terms of “market failure,” but commentators have organized them under this rubric; see, e.g., Samuelson, *supra* note 36, at 2544–46, which provides one of the most important frameworks for understanding fair use in the literature. The foremost theorist of market harm in fair use cases is Wendy Gordon, whose seminal article critiquing the Ninth Circuit’s decision in the *Betamax* case and offering a market-based test for fair use, and the subsequent expansion and clarification of her initial

example is a market in which transaction costs are too high for the parties to gain from negotiating a socially beneficial transaction, resulting in no transfer at all.¹¹⁵ Other problems that can cause markets for copyrights to fail range from effects that arise from “priceless” interests in reputation (for, say, a copyright owner whose work is parodied) or speech (for the parodist),¹¹⁶ to a failure by the market to adequately compensate for positive externalities (generated by, for example, a scholar who develops new knowledge).¹¹⁷ Whether a market has truly failed may be contested; if transaction costs are high, an additional inquiry might be required in order to decide whether this

theory by both Professor Gordon and other commentators, stand as the most complete discussion available. Gordon, *supra* note 101. *See, e.g.*, Christopher Buccafusco & Christopher Sprigman, *Valuing Intellectual Property: An Experiment*, 96 CORNELL L. REV. 1, 42–44 (2010); Wendy J. Gordon, *Fair Use Markets: On Weighing Potential License Fees*, 79 GEO. WASH. L. REV. 1814 (2011) [hereinafter Gordon, *Fair Use Markets*]; Wendy J. Gordon, *Excuse and Justification in the Law of Fair Use: Transaction Costs Have Always Been Only Part of the Story*, 50 J. COPYRIGHT SOC'Y U.S.A. 149 (2003) [hereinafter Gordon, *Excuse and Justification*]; Wendy J. Gordon, *Market Failure and Intellectual Property: A Response to Professor Lunney*, 82 B.U. L. REV. 1031 (2002) [hereinafter Gordon, *Response*]; Wendy J. Gordon, *On the Economics of Copyright, Restitution, and “Fair Use”: Systemic Versus Case-by-Case Responses to Market Failure*, 8 J.L. & INF. SCI. 7 (1997); Wendy J. Gordon, *Asymmetric Market Failure and Prisoner's Dilemma in Intellectual Property*, 17 U. DAYTON L. REV. 853 (1992); Raymond Shih Ray Ku, *Consumers and Creative Destruction: Fair Use Beyond Market Failure*, 18 BERKELEY TECH. L.J. 539 (2003); Lydia Pallas Loren, *Redefining the Market Failure Approach to Fair Use in an Era of Copyright Permission Systems*, 5 J. INTELL. PROP. L. 1 (1997); Glynn S. Lunney, Jr., *Fair Use and Market Failure: Sony Revisited*, 82 B.U. L. REV. 975 (2002); 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, 130–35 (1997). Professor Gordon's initial three-part test for determining fair use under a market-based analysis included the following elements: (1) presence of market failure; (2) transferring the use to the defendant is socially desirable; and (3) finding fair use would not substantially injure the copyright owner's incentives. Gordon, *supra* note 101, at 1614. Over time, Gordon expanded her initial analysis, and eventually decided that the third requirement should only apply in some cases. *See, e.g.*, Gordon, *Excuse and Justification, supra*, at 183–84; Gordon, *Response, supra*, at 1031–32. However, even under the original, strict form of the test, orphan works present a clear market failure that should be remedied by fair use. *See infra* notes 115–125 and accompanying text.

115. *See generally* William M. Landes & Richard A. Posner, *An Economic Analysis of Copyright Law*, 18 J. LEGAL STUD. 325 (1989); RICHARD POSNER, *ECONOMIC ANALYSIS OF LAW* (2d ed. 1977); R.H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960).

116. *See, e.g.*, Richard Posner, *When is Parody Fair Use?*, 21 J. LEGAL STUD. 67, 73–74 (1992) (“There is an obstruction when the parodied work is a target of the parodist's criticism, for it may be in the private interest of the copyright owner, but not in the social interest, to suppress criticism of the work. It is doubtful that this problem would be fully solved without an exemption from copyright protection even if the parodist could transform the social value of his work as criticism into private value and therefore compensate the copyright owner for the diminution in the latter's revenues as a result of the criticism.”).

117. *See, e.g.*, Gordon, *supra* note 101, at 1630. *See also infra* notes 118–125 and accompanying text.

actually creates such a barrier to transfer that transfers will not occur or would only occur at a socially detrimental cost.

It is difficult to imagine a more complete form of market failure than that represented by orphan works. Not only are orphans not available on the market, but a market cannot form for lack of a rights-holder with whom would-be licensees could negotiate. For orphan works, no locatable owner exists, so no matter what the level of investment in trying to negotiate, no transaction will occur. Even for a possible set of works for which some owner exists somewhere—perhaps an oblivious or uninterested heir, or many oblivious or uninterested heirs—a markedly unreasonable outlay of transaction costs could be required in order to close a transaction. This set of works also represents a serious market failure.¹¹⁸ Consequently, markets around orphan works are decidedly unlikely to form, leaving little, if any, cognizable market harm that could arise from noncommercial efforts to preserve and provide access to such works.¹¹⁹

Digitizing and making orphans accessible to library and archive patrons, then, would seem highly unlikely to invade copyright owners' markets, favoring these activities under the market-harm factor.¹²⁰ But would finding a fair use in the face of such a market failure nonetheless damage the copyright holder's incentives to create and disseminate new works in the future? While one can imagine that fear of being disassociated with one's rights in the future could disincentivize either creation of a work or, more likely, dissemination of the work to the public, this outcome also seems highly unlikely in the case of orphan works. Elsewhere in this issue, Professor Randal Picker explains how limited the likely incentive benefit of an orphan

118. *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 937 (2d Cir. 1994) (Jacobs, J., dissenting); Gordon, *supra* note 101, at 1618.

119. Compare markets for orphans, for example, to a possible market for the HathiTrust's uses of non-orphaned works in the HathiTrust Digital Library. In this case, Judge Baer found that:

Defendants offer substantial evidence that it would be prohibitively expensive to develop a market to license the use of works for search purposes, access for print-disabled individuals, or preservation purposes. Waldfoegel Decl. ¶¶ 7, 22–24 (estimating that the costs of such a license as to the works in the HDL would be in the neighborhood of \$569 million and that the potential revenue generated would not cover these costs so it was not a “commercially viable endeavor”).

Authors Guild, Inc. v. HathiTrust, 11 CV 6351 HB, 2012 WL 4808939, at *14 (S.D.N.Y. Oct. 10, 2012).

120. *See, e.g., Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 591–92 (1994); *Becker*, No. 08 Civ. 1425, 72–81; *Hofheinz v. A & E Television Networks*, 146 F. Supp. 2d 442, 448–49 (S.D.N.Y. 2001).

works licensing regime would be.¹²¹ And as explained, Section III.A, *supra*, some attributes likely to be contained in an orphan work’s “nature”—less likely to be commercially valuable, more likely created for reasons other than economic gain—may indicate that incentives are unlikely to be damaged by repurposing the prototypical orphan work.

Where there is no cognizable harm to the market for a work, we should not expect incentives to create to be damaged, and should instead fear the loss of public benefit inherent in keeping the work from being put to beneficial purposes. As Justice Stevens pointed out in *Sony v. Universal*:

“[A] use that has no demonstrable effect upon the potential market for, or the value of, the copyrighted work need not be prohibited in order to protect the author’s incentive to create. [This] would merely inhibit access to ideas without any countervailing benefit What is necessary [for a noncommercial use] is a showing by a preponderance of the evidence that *some* meaningful likelihood of future harm exists.”¹²²

“[I]nhibit[ing] access to ideas without any countervailing [public] benefit”¹²³ is precisely the type of social harm fair use seeks to prevent, and precisely the type of harm that can be prevented by library and archive digitization projects. Accordingly, the social harm created by the failure of orphan work markets is compounded by the positive externalities that would be generated by users who would “free”¹²⁴ the works and turn them to public benefit. Professor Wendy Gordon has identified two examples of positive externalities that are especially relevant here: the “significant ‘external benefits’” of teaching and scholarship, each of which enjoy significant support from libraries and archives. Gordon points out that through teaching and scholarship, “all of society benefits [by] having an educated citizenry and advances in knowledge,” though educator salaries are relatively low and the benefits of communicating ideas may be nonmonetizable.¹²⁵

Library and archive uses are especially likely to generate these and other positive externalities. Digitizing an orphan work and making it accessible makes it available to at least the patron base of a library, and in the cases of

121. Randal C. Picker, *Private Libraries and Orphan Works*, 27 BERKELEY TECH. L.J. 1259, 1280–82 (2012).

122. *Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417, 450–51 (1984).

123. *Id.* at 451.

124. Loren, *supra* note 27, at 1452–55.

125. Gordon, *supra* note 101, at 1630. *See also Sony*, 464 U.S. at 477–78 (Blackmun, J., dissenting) (“[Where] the scholar forgoes the use of a prior work, not only does his own work suffer, but the public is deprived of his contribution to knowledge. The scholar’s work, in other words, produces external benefits from which everyone profits.”).

publicly accessible archives, to everyone. Such activities create spillover educational and cultural capital that accrues to the public in the form of learning and shared cultural and historical understanding. And of course any follow-on research, scholarship, or other social output that stems from the availability of orphan works represents further positive externalities attendant to those works becoming accessible.

C. COLLECTIVE OR OTHER CENTRALLY MANAGED LICENSING

Although no market for digitizing and making orphans accessible can naturally arise, it may be possible to create a licensing market through collective licensing or similar mechanisms. There would be some advantages of a licensing market, both because of the enhanced legal certainty of licensed uses for libraries and archives, and because a well-managed system could greatly limit the risk that a later-appearing owner would be left without compensation.¹²⁶ A few such licensing systems already exist in other countries: Canada, India, Japan, South Korea, the United Kingdom, and Fiji.¹²⁷ All of these schemes require some form of diligent search for the copyright owner¹²⁸ and then delegate the right to issue licenses and set fees to a central authority. As another model, France recently adopted an extended collective license (ECL) regime that includes provision for unclaimed out-of-commerce works; after six months' notice, the right to grant licenses accrues to the central licensing authority.¹²⁹ In this and other ECL regimes, the

126. For arguments in favor of licensing for orphans, see, e.g., HARGREAVES, *supra* note 24, at 40 (recommending an extended collective license regime in the United Kingdom); Stef van Gompel & P. Bernt Hugenholtz, *The Orphan Works Problem: The Copyright Conundrum of Digitizing Large-Scale Audiovisual Archives, and How to Solve It*, 8 POPULAR COMM. 61 (2010).

127. Hansen, *supra* note 29, at 10 (describing regimes and collecting sources). Some of these schemes are extremely limited in nature. For example, the United Kingdom and Fijian laws give the Copyright Tribunal the ability to consent to “a person making a recording from a previous recording of a performance where the identity and whereabouts of a performer cannot be ascertained by reasonable inquiry.” AGNIESZKA VETULANI, *THE PROBLEM OF ORPHAN WORKS IN THE EU: AN OVERVIEW OF LEGISLATIVE SOLUTIONS AND MAIN ACTIONS IN THIS FIELD* 10 (Feb. 2008), http://ec.europa.eu/information_society/activities/digital_libraries/doc/reports_orphan/report_orphan_v2.pdf.

128. VETULANI, *supra* note 127, at 9–10. For example, the Copyright Board of Canada requires a hopeful user to “have made every reasonable effort to find the copyright owner” and the Japanese Commissioner of the Agency for Cultural Affairs may issue a license after “due diligence” fails to turn up the copyright owner. *Unlocatable Copyright Owners*, COPYRIGHT BOARD OF CANADA, <http://www.cb-cda.gc.ca/unlocatable-introuvables/brochure2-e.html> (last modified July 7, 2001); VETULANI, *supra* note 127, at 10.

129. Loi no 2012-287 du 1er mars 2012 relative à l'exploitation numérique des livres indisponibles du XXe siècle (1) [Law 2012-287 of Mar. 1, 2012 on the Digital Exploitation of Unavailable Books of the Twentieth Century], JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE [J.O.] [OFFICIAL GAZETTE OF FRANCE], Mar. 2 2012, p. 03986, *available at*

central licensing body is empowered to issue licenses for all copyright owners within its ambit, avoiding the need for a search for orphan owners.

Many of these systems cover orphans as extensions of broader schemes intended to reduce transaction costs and facilitate licensing of copyrighted works more generally¹³⁰ and are thus directed at solving a broader set of problems than only those created by orphan works. And of course, American-style fair use is not a feature of the copyright law in the countries that have employed licensing systems. This fact undoubtedly makes such systems more attractive locally. Even in the United States, centralized licensing systems may be an appropriate approach to the stark transaction costs presented by the mass digitization of non-orphaned works, at least when those works are used in mass digitization projects that do not exhibit other features of fair use—for instance, where the project’s purpose is a commercial one closely aligned with the copyright owners’ original purposes in disseminating the work.¹³¹ For example, this might be a useful approach for books that are out of print due to lack of broad demand, but that are owned by someone with whom a license could be negotiated if transaction costs could be sufficiently diminished.¹³² In such a case, using a centralized licensing scheme to lower transaction costs enough to encourage new paid uses may be an efficient approach to encouraging mass digitization.¹³³

http://www.legifrance.gouv.fr/jopdf/common/jo_pdf.jsp?numJO=0&dateJO=20120302&numTexte=1&pageDebut=03986&pageFin=03988#. See also Gillian Spraggs, *France Guillotines Copyright*, ACTION ON AUTHORS’ RIGHTS (Feb. 28, 2012), <http://blog.authorsrights.org.uk/2012/02/28/france-guillotines-copyright/> (providing a basic English translation of the law and links to commentary).

130. VETULANI, *supra* note 127, at 8–13.

131. This is an intentionally narrow example—the purpose of the use is clearly highly important to whether licensing or fair use is the most appropriate option. Theoretically available copyright owners do not necessarily mean that licensing approaches are favored over fair use. Transaction costs, endowment effects, or other economic barriers may be sufficiently high, or social benefits may be sufficiently great for digitization uses to be fair beyond this context. See, e.g., *Kelly v. Arriba Soft Corp.*, 336 F. 3d 811 (9th Cir. 2002) (reproduction of works to facilitate search engines, which require many millions or billions of unlicensed reproductions in order to function well, found to be fair use); *Perfect 10 Inc. v. Amazon.com, Inc.*, 508 F.3d 1146 (9th Cir. 2007) (same); *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1122 (D. Nev. 2006) (same). And, as discussed Part III.A, *supra*, the “nature” of the work as an orphan should be considered highly relevant in any fair use case. This Article is limited to a discussion of the digitizing of orphans by nonprofit libraries and archives; working out the parameters of when and if licensing is appropriate in other contexts is a topic for another paper. I do not take a position on that subject here.

132. There is some authority, however, suggesting that unavailability on the market should weigh in favor of fair use. See *supra* notes 68–72 and accompanying text.

133. Symposium, *Collective Management of Copyright: Solution or Sacrifice?*, 34 COLUM. J.L. & ARTS 589 (2011). This is not to say that collective licensing is necessarily the best solution

For the library and archive uses of true orphan works discussed here, however, fair use is likely to prove a far more socially beneficial approach.¹³⁴ The apparent benefits of certainty for the licensee and compensation for a later-appearing owner are likely to be far outweighed by the costs of a collective license system.

There are at least three sets of costs attendant to a centralized scheme for licensing orphans: first are the transaction costs required to bring licensees and copyright owners together and create a fair bargain; second are the administrative costs required to develop and manage the licensing regime; and third are the licensing fees. Because orphans, by definition, represent the failure of traditional licensing systems, some additional investment above and beyond that required by a naturally arising and functioning market would be required. This effect is exacerbated by the fact that a large portion of the costs of such a licensing system would likely constitute a substantial inefficient misallocation of resources, because most possible owners either will not exist or will be impossible to find at any price.

Considering each set of costs in turn illuminates the social harm that recognizing fair use would prevent. With regard to transaction costs, for true orphans, no amount of investment can result in a meaningful transaction, because the licensor is simply unavailable to complete it. While a generalized collective license could certainly be created, the possibility of such a license representing an actual connection between licensee and licensor is vanishingly small. No reasonable amount of investment in transaction costs is likely to achieve a match, even for works that could theoretically be “cleared” with enough expenditure. For true orphans, no amount of investment will produce a match, at all.

even for this situation. To be viable, any such system requires, at a minimum, the infusion of administrative resources, robust and effective oversight mechanisms, and safeguards to ensure that license prices are fair for both copyright holders and licensees and that resources are not misallocated. Further study is required, and full discussion is outside the scope of this Article. Commentators offer competing analyses of collective licensing approaches for different types of works. WILLIAM W. FISHER III, PROMISES TO KEEP: TECHNOLOGY, LAW, AND THE FUTURE OF ENTERTAINMENT 199–258 (2004); Jonathan Band, *The Book Rights Registry in the Google Book Settlement*, 34 COLUM. J.L. & ARTS 671 (2011); Fred von Lohmann, *Monetizing File-Sharing: Collective Licensing Good, ISP Tax Bad*, Electronic Frontier Foundation (Mar. 20, 2008), <https://www.eff.org/deeplinks/2008/03/monetizing-file-sharing-collective-licensing-good-isp-tax-bad>.

134. It is perhaps instructive that none of the countries that have adopted licensing systems have United-States-style fair use. See Peter Jaszi, *Public Interest Exceptions in Copyright: A Comparative and International Perspective* (presented at Correcting Course: Rebalancing Copyright for Libraries in the National and International Arenas), available at http://correctingcourse.columbia.edu/paper_jaszi.pdf.

Administrative costs present a similar issue: where the likelihood that the system can effectively transfer payment from licensees to licensors is as low as it is with orphans, the costs inherent in setting up and maintaining a licensing structure will nearly always be too high for efficiency to be achieved. For example, the Canadian system requires what is likely a quite costly *ex ante* review of requested orphan work uses.¹³⁵ Under this system, the Copyright Board of Canada issued only 216 orphan work licenses between 1990 and 2007.¹³⁶ (The rate did increase from an average of four to twenty-three licenses issued per year over that time period.¹³⁷) While administrative costs will certainly vary by system, any system would have to be sufficiently robust to reliably offer transactions, even though the transactions are unlikely to pay copyright owners who have gone missing. Accordingly, both transaction costs and administrative costs represent potentially substantial misallocations of resources away from making orphans accessible, developing systems to prevent orphaning in the future, and other social benefits.

Finally, most licensing fees themselves would be misallocated from the putative licensee to the licensing system or another beneficiary. No owner exists to claim them, so even the lower fees that can be created by lowering transaction costs through centralized licenses are socially harmful.¹³⁸ As noted, the possibility of a collective license representing actual connections between licensors and licensees is very low where the owner is unlocatable. None of the existing license regimes present a satisfactory solution to this problem. For instance, the Nordic ECL systems delegate decisions about fund distribution to the relevant Copyright Management Organization,¹³⁹ and under the Canadian system the Copyright Board of Canada or collective society it tasks with distributing funds may use uncollected funds as is seen fit.¹⁴⁰ While clear policies for disbursement and robust oversight procedures

135. BERNT HUGENHOLTZ ET AL., *THE RECASTING OF COPYRIGHT AND RELATED RIGHTS FOR THE KNOWLEDGE ECONOMY* 187 (2006), available at http://ec.europa.eu/internal_market/copyright/docs/studies/etd2005imd195recast_report_2006.pdf.

136. VETULANI, *supra* note 127, at 10.

137. *Id.*

138. For a fuller discussion of this issue, see Picker, *supra* note 121.

139. JOHAN AXHAMN & LUCIE GUIBAULT, *CROSS-BORDER EXTENDED COLLECTIVE LICENSING: A SOLUTION TO ONLINE DISSEMINATION OF EUROPE'S CULTURAL HERITAGE?* 41 (2011), available at http://www.ivir.nl/publicaties/guibaut/ECL_Europeana_final_report_092011.pdf.

140. The Copyright Board may set license terms, royalties, and disbursement. Copyright Act, R.S.C. 1985, c. C-42, s. 77 (Can.). The Board typically pays collected royalties to a copyright collective society "that would normally represent the unlocatable copyright owner." The collective society, in turn, pays the royalties out to the owner, should one

could certainly help central authorities manage and disburse funds as effectively as possible, there is no way to address the fundamental inefficiency inherent in allocating purported license fees away from the copyright owner, and certainly no reason to think that a central authority would be more effective at picking socially beneficial disbursements than those who would make fair use of orphan works.¹⁴¹

There is one further cautionary note on licensing markets for orphans: the danger of undermining the application of fair use by, albeit artificially, creating a “market” for orphan works that could then be harmed by otherwise fair uses of the works—even though the overall social harm is great and the benefit to actual authors minimal. This “circularity” problem with the market-harm factor is familiar to both courts and commentators, who have wrestled with it for years.¹⁴² Licensing mechanisms have proven beguiling to courts considering market harm in the past,¹⁴³ though recent case law has begun to clarify the issue, at least for transformative uses of works.¹⁴⁴ Circularity is of particular concern with regard to the use of orphan works for educational or other socially beneficial purposes, because the social costs

appear within five years after the license expires; if no owner comes forward, fees are used as the collective society sees fit. *Unlocatable Copyright Owners*, *supra* note 128.

141. Indeed, the U.S. Copyright Office explicitly declined to recommend an approach like the Canadian system for this reason, COPYRIGHT OFFICE, *supra* note 8, at 83, 114, and the subsequent legislation proposed in Congress adopted a remedies-limitation approach instead. Orphan Works Act of 2006, H.R. 5439, 109th Cong., at 9 (2006); Orphan Works Act of 2008, H.R. 5889, 110th Cong., at 14 (2008); Shawn Bentley Orphan Works Act of 2008, S. 2913, 110th Cong., at 15 (2008).

142. I cannot do justice to the discussion in the space available here, but see, e.g., *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 614–15 (2d Cir. 2006) (“[A] copyright holder cannot prevent others from entering fair use markets merely ‘by developing or licensing a market . . .’”); *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 937 (2d Cir. 1994) (Jacobs, J., dissenting); James Gibson, *Risk Aversion and Rights Accretion in Intellectual Property Law*, 116 YALE L.J. 882, 947–51 (2007); Wendy J. Gordon, *The ‘Why’ of Markets: Fair Use and Circularity*, 116 YALE L.J. POCKET PART 371 (2007) (responding to Gibson); Gordon, *Fair Use Markets*, *supra* note 101 at 1826–34 (discussing the controversy more generally); Loren, *supra* note 114, at 38–39; Samuelson, *supra* note 36, at 2620 (discussing the controversy).

143. *Princeton University Press v. Michigan Document Services*, 99 F.3d 1381 (6th Cir. 1996), is a well-known example. In this case, a number of copy shops were persuaded to pay license fees for making educational course packets. Princeton University Press then sued a hold-out copy shop, which claimed fair use. Though others had operated without licenses for years, the court accepted that the licenses constituted market harmed by unlicensed use, and found against fair use. In dissent, Judge Merritt argued that this reasoning was problematic. *Id.* at 1397 (Merritt, J., dissenting) (“If the publishers have no right to the fee in many of the instances in which they are collecting it, we should not validate that practice by now using the income derived from it to justify further imposition of fees.”).

144. See, e.g., *Bill Graham*, 448 F.3d at 614–15.

inherent in a licensing system are so steep. It would be a poor outcome for the public and for copyright if a system that facilitated as few uses of orphans as the Canadian system does was then seen as a relevant market to be harmed by otherwise beneficial uses, limiting the application of fair use.

Overall, given the lack of benefit and the social cost of creating and maintaining them, licensing systems seem inadvisable where fair use supports making orphans accessible. Thus, courts should weigh the market-harm factor accordingly. For uses by nonprofit libraries and archives, which are both especially sensitive to cost and especially likely to generate positive externalities through their efforts to digitize orphans and make them available, the market-harm factor should weigh especially heavily towards fair use.

IV. NONCOMMERCIAL USE OF AND ACCESS TO ORPHAN WORKS VIA LIBRARIES AND ARCHIVES PROMOTES THE PROGRESS OF SCIENCE

A. THE PURPOSE AND CHARACTER OF NONPROFIT LIBRARY AND ARCHIVE USES OF ORPHANS FAVORS FAIR USE

The first and often most important question considered in fair use cases is whether the purpose of the defendant's use furthers core societal goals—such as learning, access to information, freedom of speech and expression, and innovation¹⁴⁵—or whether such use simply interferes with the rights of the copyright holder to exploit her rights in the work. Section 107 directs courts to consider “whether such use is of a commercial nature or is for nonprofit educational purposes.”¹⁴⁶ Section 107 also provides a list of six favored uses: “criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, [and] research.”¹⁴⁷ The purpose inquiry helps courts assess whether allowing a defendant's use would better fulfill copyright's purpose to promote Progress through the “creation and spread of knowledge” than would giving the copyright holder veto power over the use. Accordingly, the purpose factor constitutes a core component of fair use cases and, together with the market-harm factor, is one of the most influential factors in the analysis.¹⁴⁸

145. Samuelson, *supra* note 36, at 2544–46 (organizing fair use cases into “policy-relevant clusters,” including these).

146. 17 U.S.C. § 107(1) (2010).

147. *Id.*

148. Beebe, *supra* note 49, at 597–610; Sag *supra* note 98, at 55–61.

Historically, there have been few cases analyzing libraries' and archives' purposes in copying works. As this Article was going to press, however, Judge Baer's decision in *Authors Guild v. HathiTrust* was published. This decision squarely addresses several purposes—full-text search, preservation, and access for users with print disabilities—for which the HathiTrust libraries are using the Google-scanned books in the Google Book Search corpus.¹⁴⁹ As such, it sheds new light on how at least one federal district court views library purposes for mass digitization, and I have incorporated some of Judge Baer's findings into this analysis.¹⁵⁰

The “creation and spread of knowledge” is central to the mission of nonprofit libraries and archives. This mission includes both preserving the knowledge contained within books, films, papers, and other artifacts, and spreading that knowledge.¹⁵¹ Consequently, when libraries and archives digitize orphan works and make them electronically available to patrons or to the broader public, they perform a variety of functions that are of a purpose and character likely to weigh steeply in favor of fair use under current doctrine. As discussed below, these libraries also perform functions for which there is little, or mixed, caselaw but that are so strongly in the public interest that they should weigh in favor of fair use nonetheless. Where the equities are close in weight, library and archives' salutary social functions, together with the lack of probable harm that would arise from making orphans available, should tip such functions toward fair use.

These functions can be divided into two sets: first-order uses and second-order uses.¹⁵² First-order uses are those undertaken directly by the library or archive. These uses include the initial digitization of the work, copying done for preservation, copying done to create indexes, catalogs, or other mechanisms that facilitate searching for and within the digitized works,

149. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *2 (S.D.N.Y. Oct. 10, 2012).

150. Judge Baer's decision does not address orphan works. This is because the HathiTrust's nascent Orphan Works Program was stopped upon commencement of the litigation; Judge Baer thus held that the issue was not ripe for adjudication. *Id.* at *7–8. The decision instead addresses uses the HathiTrust made of books with known authors. *Id.*

151. *About the Internet Archive*, INTERNET ARCHIVE, <http://archive.org/about/about.php> (“Libraries exist to preserve society's cultural artifacts and to provide access to them.”); AMERICAN LIBRARY ASSOCIATION & ASSOCIATION OF RESEARCH LIBRARIES, THE ALA AND ARL POSITION ON ACCESS AND DIGITAL PRESERVATION: A RESPONSE TO THE SECTION 108 STUDY GROUP (Nov. 9, 2006), *available at* <http://www.ala.org/advocacy/sites/ala.org/advocacy/files/content/copyright/dmca/section108/sec108study/108ppfinal.pdf> (“The mission of libraries is to preserve and provide access to information, regardless of format.”).

152. Thank you to Dave Hansen for suggesting this helpful terminology.

and copying done to place the works in historical and cultural context. Second-order uses are those supported by libraries' and archives' direct efforts but undertaken by others, usually patrons or members of the public, such as reading, viewing, conducting research, and producing scholarship or other secondary works.

To begin with first-order uses, most activities traditionally undertaken by libraries and archives further "nonprofit educational purposes"¹⁵³ and provide other "public benefit[s],"¹⁵⁴ putting them within the core of purposes that weigh in favor of fair use. Both Congress and the courts have recognized the public benefit inherent in preserving cultural materials. For example, Congress recognized an important public benefit in preserving films recorded on volatile nitrate stock, which it explicitly called out as a fair use.¹⁵⁵ In *Sundeman v. Seajay Society*,¹⁵⁶ the Fourth Circuit held that the complete reproduction of an unpublished, "fragile, seventy year old original manuscript" by a library was fair use, and "unquestionably served the 'public benefit' and the 'development of art.'"¹⁵⁷ And in *Authors Guild v. HathiTrust*, Judge Baer held that preservation of books via the mass digitization of print copies was a fair use.¹⁵⁸

153. 17 U.S.C. § 107(1) (2010).

154. *Am. Geophysical Union v. Texaco, Inc.*, 802 F. Supp. 1, 27 (S.D.N.Y. 1992), *aff'd*, 60 F.3d 913 (2d Cir. 1994).

155. H.R. Rep. No. 94-1476, at 67 (1976) ("The efforts of the Library of Congress, the American Film Institute, and other organizations to rescue and preserve this irreplaceable contribution to our cultural life are to be applauded, and the making of duplicate copies for purposes of archival preservation certainly falls within the scope of 'fair use.'"). More generally, of course, it also decided that preservation activities by qualifying libraries and archives justified a specific exception in § 108 of the Copyright Act. 17 U.S.C. § 108(b) (2010) (covering some limited copying for preservation purposes of unpublished works); 17 U.S.C. § 108(c) (2010) (excepting reproduction in some cases where the existing format in which the work is stored has become obsolete). In his recent article, *The Impact of Substantial Compliance with Copyright Exceptions on Fair Use*, Jonathan Band makes precisely this point, asking "[w]hat better indication could there be of the kinds of other purposes favored by Congress for fair use than the specific exceptions contained in the Copyright Act?" *See* Band, *supra* note 18, at 6 (exploring the relationship between the specific exceptions and fair use).

156. 142 F.3d 194 (4th Cir. 1998). As noted *supra* note 18, Congress' willingness to create exceptions for library preservation and other purposes most assuredly should not be considered evidence against those libraries' abilities to claim fair use, contrary to the Authors Guild's argument in Complaint, *Authors Guild, Inc. v. HathiTrust*, No. 11 Civ. 6351. Jonathan Band's analysis in Band, *see supra* note 18, is much more sensible.

157. *Sundeman*, 142 F.3d at 203.

158. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *10 n.19 (S.D.N.Y. Oct. 10, 2012). Note that Judge Baer does state that, in his view, "[t]he argument that preservation on its own is a transformative use is not strong." *Id.* at 15. However, he saw the "noncommercial nature of the use" as an important distinction from previous cases,

Libraries and archives seek to preserve a wide variety of orphan works: books, photographs, sound recordings, visual works,¹⁵⁹ rapidly disintegrating films,¹⁶⁰ and websites,¹⁶¹ among many other materials. Preventing the loss of these cultural resources for future generations of readers, researchers, and learners clearly provides an important public benefit and serves a socially valuable purpose. With regard to orphan works specifically, preservation is likely of even greater value, as with no owner to preserve them or give permission for others to do so, orphans are at particular risk of being lost.

Libraries and archives also routinely promote learning and access to information by organizing materials, making them searchable, providing direct access to them, and offering productive enhancements such as curation,¹⁶² direct contextualization, or commentary.¹⁶³ Organizing materials and facilitating information access through full-text searchable databases of these works, as search engines have done with images on websites, is likely to be considered transformative under current case law.¹⁶⁴ In any case, this use is clearly a productive use that benefits the public.¹⁶⁵ This prediction has been borne out by the *HathiTrust* decision, in which Judge Baer held that digitizing

especially *Texaco*, and cited the House Report, discussed *supra* note 155, in support of this analysis. *Id.*

159. Response of Library Copyright Alliance to Orphan Works Notice of Inquiry, Mar. 25, 2005, Comment OW0658 at 3–11, *available at* <http://www.copyright.gov/orphan/comments/OW0658-LCA.pdf>.

160. *See* Schwartz, *supra* note 82, at slide 6; *Access to Orphan Films*, Response of Center for the Study of the Public Domain to Orphan Works Notice of Inquiry, Mar. 2005, Comment OW0596 at 1–4, *available at* <http://www.copyright.gov/orphan/comments/OW0596-CPD.pdf>.

161. *About the Wayback Machine*, INTERNET ARCHIVE, <http://archive.org/web/web.php> (last visited Nov. 5, 2012).

162. LIBRARY OF CONGRESS, *Mission*, *supra* note 2; *About the Grateful Dead Archive*, *supra* note 14.

163. *See, e.g., Patriotic Melodies*, *supra* note 5.

164. *See* *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1164–67 (9th Cir. 2007); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818–20 (9th Cir. 2003); *Field v. Google*, 412 F. Supp. 2d 1106, 1118–19 (D. Nev. 2006).

165. *See id.* Note that the works in question in each of these cases were publicly available on the Internet at the direction of the copyright holder, while libraries and archives would often be making the works available themselves. However, the lack of harm to missing owners of orphans is at least arguably analogous to the lack of harm to owners in these cases. *See also* *New York Times, Inc. v. Roxbury Data Interface, Inc.*, 434 F. Supp. 217 (D.N.J. 1977) (holding that preparing a personal names index to a set of New York Times indices was fair use and stating, “[The index] appears to have the potential to save researchers a considerable amount of time and, thus, facilitates the public interest in the dissemination of information.”).

to create “superior search capabilities” was indeed transformative, and was fair use.¹⁶⁶

Providing access to works—both in their original forms and in aggregated forms that allow for text-mining or other non-display uses¹⁶⁷—helps fulfill copyright’s most basic goal to “promote . . . Progress” through the dissemination of knowledge.¹⁶⁸ And adding contextualization and commentary to works repurposes them in a “productive” or “transformative” manner held by courts to be at the core of fair use’s support for free expression.¹⁶⁹ Each of these public benefits is again more valuable for orphans, which may otherwise languish in obscurity, and for which there is no existing original market to harm.

The first-order uses directly made by libraries and archives, of course, are also intended to support second-order educational, research, scholarly, critical, and other salutary uses made by library and archive patrons, including consumption for private learning. It is through these uses that copyright’s dissemination goals are actually met. In organizing research, scholarship, and commentary into a cluster of “authorial” uses, for example, Professor Pamela Samuelson notes that such uses are ones “that copyright law welcomes in order to promote the ongoing creation and dissemination of new knowledge, thereby fulfilling the constitutional purpose of copyright law.”¹⁷⁰ Accordingly, these should comfortably be considered fair use, and courts tend to hold that they are.¹⁷¹ In the *HathiTrust* case, for example, Judge Baer specifically tied the defendant libraries’ digitization program’s support of learning by users with print disabilities to copyright’s goals, citing these users’ “unprecedented ability . . . to have an equal opportunity to compete with their sighted

166. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *11 (for first point), *14 (for holding) (S.D.N.Y. Oct. 10, 2012).

167. See Matthew Sag, *Orphan Works as Grist for the Data Mill*, 27 BERKELEY TECH. L.J. 1503 (2012).

168. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 575 (1994) (quoting U.S. CONST., art. 1, § 8, cl. 8); EU ORPHAN WORKS DIRECTIVE, *supra* note 24; HARGREAVES, *supra* note 24, at 39; COMITÉ DES SAGES, *supra* note 24.

169. See, e.g., *Campbell*, 510 U.S. at 586–87; *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 609 (2d Cir. 2006).

170. Samuelson, *supra* note 36, at 2545.

171. See *Sundeman v. Seajay Society*, 142 F.3d 194 (4th Cir. 1998) (fair to make full copy and use excerpts for scholarly commentary); *Time, Inc. v. Bernard Geis Assocs.*, 293 F. Supp. 130 (S.D.N.Y. 1968) (fair use to publish sketches that recreate frames from the Zapruder Kennedy assassination film to illustrate a book offering alternative theories about the assassination); see also Samuelson, *supra* note 36, at 2546–78 (discussing uses that promote free expression and new authorship).

peers.”¹⁷² And while there is a dearth of authority regarding private, noncommercial—or “personal”—uses, there is some evidence that Congress thought such uses should be fair,¹⁷³ and some courts have held them to be so.¹⁷⁴ Where these uses do not damage a copyright holder’s legitimate market, commentators have persuasively argued that they should be fair in order to protect legitimate and customary expectations in privacy,¹⁷⁵ self-expression,¹⁷⁶ and autonomy.¹⁷⁷

Although educational and research uses are statutorily favored, the case law dealing with large-scale, institutional uses has less predictably found fair use than might be expected in light of copyright’s purpose to promote learning.¹⁷⁸ It remains to be seen whether the *HathiTrust* decision will have a broader effect, but it is a promising contrary example to this trend, given its close factual ties to libraries’ actual practices and goals. In any case, when the work to be used is an orphan, courts need not worry about undermining publisher incentives to make works available, so the purpose factor should weigh more heavily in favor of fair use. In the principal research or educational use cases where fair use defenses have failed, orphan works were

172. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *14 (S.D.N.Y. Oct. 10, 2012).

173. *See* H. Rep. 94-1476, at 74 (1976) (stating personal uses for education and research purposes may be fair); S. Rep. 94-473, at 63–66 (1975) (same). When considering personal uses, Congress was likely following the lead of other countries, many of which have explicit personal use exceptions built into their copyright laws. *See, e.g.*, *BMG Canada Inc. v. Doe*, [2004] F.C. 488 (Can. Fed. Ct.) (“[D]ownloading a song for personal use does not amount to infringement.”).

174. *See, e.g.*, *Sony Corp of America v. Universal City Studios, Inc.*, 464 U.S. 417, 454–55 (1984) (holding private noncommercial time-shift copying of television programs fair use); *Recording Indus. Ass’n Am. v. Diamond Multimedia Sys. Inc.*, 180 F.3d 1072, 1079 (9th Cir. 1999) (characterizing place-shifting as “paradigmatic noncommercial personal use”). *But see* *A&M Records, Inc. v. Napster, Inc.*, 239 F. 3d 1004, 1019 (9th Cir. 2001) (distinguishing *Sony* and *Diamond* because “shifting in these cases did not also simultaneously involve distribution of the copyrighted material to the general public”); *Metro-Goldwyn Mayer Studios, Inc. v. Grokster, Ltd.*, 545 U.S. 913, 931–32 (2005); *see also* *Sega Enterp. Ltd. v. MAPHIA*, 857 F. Supp. 679 (N.D. Cal. 1994) (rejecting *Sony* private use defense by commercial bulletin board service held as indirect infringer for facilitating up and downloading of Sega games by its users); *Playboy Enterp., Inc. v. Frena*, 839 F. Supp. 1552 (M.D. Fla. 1993) (rejecting *Sony* private use defense by commercial bulletin board service held as infringer for facilitating up- and downloading of Playboy bunny pictures by its users).

175. *See* Julie E. Cohen, *The Place of the User in Copyright Law*, 74 *FORDHAM L. REV.* 347, 349 (2005); Jessica Litman, *Lawful Personal Uses*, 85 *TEXAS L. REV.* 1871, 1915 (2007).

176. *See* Joseph P. Liu, *Copyright Law’s Theory of the Consumer*, 44 *B.C. L. REV.* 397, 415–20 (2003).

177. *See* Liu, *supra* note 176, at 406–11.

178. Samuelson, *supra* note 36, at 2545 (“Sharply divergent views on fair use exist in the educational and research use case law, and [these uses are least] predictable.”).

not at issue, nor were library or archive programs directly considered. Publishers won challenges to for-profit copy shops selling coursepacks of articles and book chapters without permission,¹⁷⁹ and to the for-profit company Texaco's scientists making archival copies of journal articles that publishers were willing to license.¹⁸⁰ Yet, a fair use defense succeeded in a case challenging the National Institutes of Health and the National Library of Medicine for systematically photocopying scientific articles for the benefit of government medical and scientific researchers.¹⁸¹ And most analogously and recently prior to the *HathiTrust* decision, the trial court in *Cambridge University Press v. Becker* found that it was fair use for Georgia State University—an actual nonprofit educational institution, rather than a for-profit copyshop—to copy “decidedly small” excerpts from books on behalf of faculty who wished to use such excerpts in teaching in the vast majority of challenged instances.¹⁸² The court further found that the arguments for fair use were strongest where an appropriate license at a reasonable price was not available. In the infringing instances, on the other hand, appropriate licenses were available.

Libraries' and archives' nonprofit digitization of orphans are likely to be much more analogous to Georgia State's or the NIH's activities, and much less like the copyshops' or Texaco's. Nonprofit libraries and archives' purposes—in short, to create public benefits—are more similar to those of government research agencies or nonprofit educational institutions than for-profit companies. In addition, like the instances of copying the court in the Georgia State case found most likely to be fair, there are no licenses available for orphans. In *Unbundling Fair Uses*, Professor Samuelson explains that institutional copying has been more controversial than “individual student or teacher copying,” in part due to the fact that “publishers of educational materials have understandably worried that very liberal fair use rules would undermine sales of books, journals, and other materials and the development of new licensing markets.”¹⁸³ Where the work in question has been orphaned, however, concerns about invading authors' or publishers' markets for educational and research materials are unlikely to arise.

179. *See* *Basic Books, Inc. v. Kinkos Graphics Corp.*, 758 F. Supp. 1522 (S.D.N.Y. 1991); *Princeton Univ. Press v. Michigan Document Services, Inc.*, 99 F.3d 1381 (6th Cir. 1996).

180. *American Geophysical Union v. Texaco, Inc.*, 60 F.3d 913 (2d Cir. 1994).

181. *Williams & Wilkins v. United States*, 487 F.2d 1345 (Ct. Cl. 1973), *aff'd by an equally divided Court*, 420 U.S. 376 (1975).

182. *Cambridge Univ. Press v. Becker*, 863 F. Supp. 2d 1190, 1232 (N.D. Ga. 2012)

183. Samuelson, *supra* note 36, at 2583–84.

In addition, some courts consider it transformative or otherwise productive to employ works for a purpose “plainly different from the original purpose for which they were created.”¹⁸⁴ As noted above, the purpose for which many orphans in library and archive collections were created is likely to be very different from the purposes to which libraries and archives wish to put them. Recall examples of works discussed above. These include items such as minutes from the SNCC May 1964 planning meeting,¹⁸⁵ once a record of a group’s organizational decisions and now a key document for understanding the historical features of the civil rights movement, or the film “U.S. Navy of 1915,” made to “convince isolationists of the importance of building a strong American Navy,” now also an historical document that would most likely be used for understanding early-twentieth-century propaganda rather than to persuade citizens of the Navy’s importance.¹⁸⁶ Judge Baer applied the concept directly in the *HathiTrust* case, noting that one new “purpose is superior search capabilities rather than actual access to copyrighted material”¹⁸⁷ and another is to “provide access to print-disabled individuals on an equal footing with sighted individuals.”¹⁸⁸ He also approvingly cited using works for “new methods of academic inquiry,” such as humanities research based on text mining.¹⁸⁹

Finally, the organization of and access to orphan works that library and archives can provide may enhance the value of those works directly—beyond the positive externalities generally expected from educational, scholarly, and other productive uses—by making them more available to discovery by those who would find new uses for them. As described above, this availability has obvious societal value, but it may have private value as well. In the rare case, owners—such as the oblivious or uncaring heirs posited in Section III.B, *supra*—may become aware, reappear, and take charge of the work. In the more likely case, other follow-on users may give the works new life.

Overall, the purposes to which libraries and archives are likely to put orphan works in their collections are likely to strongly favor fair use.

184. *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F. 3d 605, 609 (2d Cir. 2006). This analysis is also associated with the full-text search cases. *See A.V. v. iParadigms, LLC*, 562 F.3d 630, 640 (4th Cir. 2009); *Perfect 10 v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007); *Kelly v. Arriba Soft Corp.*, 336 F. 3d 811, 819 (9th Cir. 2003).

185. *See supra* note 91 and accompanying text.

186. *See supra* note 90 and accompanying text.

187. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *11 (S.D.N.Y. Oct. 10, 2012).

188. *Id.*

189. *Id.* (citing an amicus curiae brief by digital humanities scholars describing text-mining-based historical research).

B. LIBRARY AND ARCHIVE DIGITIZATION OF ENTIRE ORPHAN WORKS
OFTEN SHOULD NOT WEIGH AGAINST FAIR USE UNDER THE
AMOUNT-AND-SUBSTANTIALITY FACTOR

The final factor to be considered is “the amount and substantiality of the portion used in relation to the copyrighted work as a whole.”¹⁹⁰ In undertaking this analysis, courts consider both the actual quantitative amount used and the qualitative significance of it to the copyrighted work.¹⁹¹ Broadly speaking, the more a defendant uses of a copyrighted work, the less likely the use is to be deemed fair, and the less the defendant uses, the more likely the use will be deemed fair.¹⁹² However, digitizing and displaying entire orphan works for a variety of non-profit library and archive purposes should most often be considered neutral in a fair use analysis that reflects both the Supreme Court’s guidance to connect the amount-and-substantiality factor to other factors and the approaches courts have taken in analogous cases.

While commentators have sometimes stated that it is generally not fair use to reproduce the entirety of a work,¹⁹³ this suggestion is an oversimplified description of the case law that can lead to misunderstanding. Rather, as stated by the Supreme Court in *Campbell v. Acuff-Rose*, “the extent of the permissible copying varies with the purpose and character of the use,”¹⁹⁴ and the most salient question is thus whether the amount used is “reasonable in relation to the purpose” for which the work is being used.¹⁹⁵ Accordingly, courts have found copying entire works justified for a variety of purposes, from “time shifting” television shows,¹⁹⁶ to creating search engine indexes,¹⁹⁷

190. 17 U.S.C. § 107(3) (2010).

191. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 586–87 (1994); *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 564–66 (1985).

192. *Campbell*, 510 U.S. at 574, 587; *Harper & Row*, 471 U.S. at 561, 564–65.

193. *See, e.g.*, NIMMER, *supra* note 54, § 13.05[A][3] (“In general, it does not constitute a fair use if the entire work is reproduced.”).

194. *Campbell*, 510 U.S. at 586–87.

195. *Id.* at 586. For an example of the detailed application of this principle to specific facts, see *Warner Brothers Entertainment, Inc. v. RDR Books*, 575 F. Supp. 2d 513 (S.D.N.Y. 2008), in which the court found that because the use in question was only “slightly transformative . . . the amount and substantiality of the portion copied . . . weigh[ed] more heavily against a finding of fair use.” 575 F. Supp. 2d at 548–49. *RDR Books* is somewhat unusual in its focus on the amount-and-substantiality factor. *See* Beebe, *supra* note 49, at 615–16.

196. *Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417 (1984) (holding that “time shifting” was a fair use, but not explicitly analyzing the amount taken).

197. *Perfect 10 v. Amazon.com, Inc.*, 508 F.3d 1146, 1166 (9th Cir. 2007) (concluding “that the significantly transformative nature of Google’s search engine, particularly in light of its public benefit, outweighs Google’s superseding and commercial uses of the thumbnails in this case”); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 820–22 (9th Cir. 2003) (concluding

to building plagiarism detection programs for student papers,¹⁹⁸ to reverse engineering software,¹⁹⁹ to copying concert posters in order to display them on an historical timeline.²⁰⁰ In each of these cases, because using the entire work was reasonable or necessary to accomplishing a favorable purpose, the amount-and-substantiality factor was neutral in the analysis.

Properly considering the purpose and the amount factors together is especially important for library and archive digitization projects, which, by necessity, are very likely to reproduce entire works. In many, if not most, instances when libraries and archives digitize entire works, this digitization will be reasonable because it fulfills appropriate educational, scholarly, access to knowledge, and other socially beneficial purposes.

A few examples illustrate. In general, works cannot effectively be preserved in digital form unless they are digitized in their entirety, or as close to it as possible. The quality and accuracy of research and scholarship done using digitized materials could be compromised if scholars were limited to incomplete versions of works. For example, a historian comparing contemporary accounts of an event, a literature scholar analyzing a novel, and a scientist researching the development of a scientific theory over time each require access to entire works in order to effectively develop facts and draw conclusions.²⁰¹ Research that depends on the computational analysis of large numbers of works—for example, the “non-consumptive,”²⁰² “non-expressive,”²⁰³ or “non-display” text-mining research developing as a methodology across a variety of disciplines—could not as effectively be done with partial digital copies.²⁰⁴ Similarly, educational purposes often require that

“that Arriba’s use of Kelly’s images as thumbnails in its search engine is a fair use”); *Field v. Google*, 412 F. Supp. 2d 1106, 1120–21 (D. Nev. 2006) (noting that “like the fair uses in *Sony* and *Kelly*, Google’s use of entire Web pages in its Cached links serves multiple transformative and socially valuable purposes”).

198. *A.V. v. iParadigms, LLC*, 562 F.3d 630, 642 (4th Cir. 2009).

199. *Sega Enters. v. Accolade, Inc.*, 977 F.2d 1510, 1514, 1526–28 (9th Cir. 1992); *Sony v. Connectix*, 203 F.3d 596, 606 (9th Cir. 2000).

200. *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 613 (2d Cir. 2006).

201. While scholars will make do with limited information if necessary, this approach can also limit the strength of their analyses and thus the value of their research—a price that fair use doctrine should not, and appears not to, require the public to pay. *See, e.g., iParadigms*, 562 F.3d at 642; *Perfect 10*, 508 F.3d at 1167–68; *Kelly*, 336 F.3d at 820–22; *Field*, 412 F. Supp. 2d at 1120–21.

202. Amended Settlement Agreement, *Authors Guild, Inc. v. Google, Inc.*, Case No. 05 CV 8136 (DC) § 1.93 (S.D.N.Y. Nov. 13, 2009) (settlement as originally proposed), available at http://thepublicindex.org/docs/amended_settlement/amended_settlement.pdf.

203. *See Sag*, supra note 167, at 1535–42.

204. For example, Matt Jockers uses text mining to analyze the differences among British, American, and Irish works’ references to American Indians. *See* Matthew Jockers,

substantial portions or entire works be accessible in order for learners to gain direct experience with the materials, make meaningful comparisons between sources, and the like.²⁰⁵ And regardless of whether each individual end user requires access to entire works, a library or archive is likely to need initially to copy entire works in order to support the varying needs of different end users.²⁰⁶ The *HathiTrust* case again illustrates. Judge Baer affirmed the need to make entire copies for both searching and supporting end user needs—in this case, the needs of users with print disabilities.²⁰⁷

In addition to the connection between the amount-and-substantiality factor and the purpose factor, courts also connect the amount-and-substantiality factor to the market factor. The overarching question is whether the defendant's use creates a potential market substitution for the copyrighted work,²⁰⁸ where relevant, using more of the copyrighted work is more likely to create a substitution effect. For orphan works, however, licenses are entirely unavailable—there is no functioning market in which a substitution effect can be felt, at all.²⁰⁹ Consequently, how much of a work is

Stanford University Literary Lab, Computing and Visualizing the 19th Century Literary Genome (Apr. 13, 2012) (slides available at <http://www.law.berkeley.edu/files/Jockers.pdf>). His conclusions would have little meaning without some confidence that the data is drawn from a reasonably large and representative collection of works. See Hugh Craig, *Stylistic Analysis and Authorship Studies*, in A COMPANION TO DIGITAL HUMANITIES (Susan Schreibman, Ray Siemens & John Unsworth, eds., 2004) (noting the value of large representative samples for literary analyses); see also Nancy Ide, *Preparation of Linguistic Corpora*, in A COMPANION TO DIGITAL HUMANITIES, *supra* (explaining the importance of developing a representative sample of works for linguistic analysis).

205. See, e.g., TERRI BAYS ET AL., CODE OF BEST PRACTICES IN FAIR USE FOR OPENCOURSEWARE 11 (2009), <http://www.centerforsocialmedia.org/sites/default/files/10-305-OCW-Oct29.pdf> (giving the example of copyrighted materials that “are literally the subject of the course, rather than useful or incidental adjuncts to it or even examples or illustrations of the subject matter”); Arthur Marwick, *Primary Sources: Handle with Care*, in SOURCES AND METHODS FOR FAMILY AND COMMUNITY HISTORIANS: A HANDBOOK (Michael Drake & Ruth Finnegan eds., 1997).

206. For example, one patron may need to review only a chapter, or portion of a chapter, within a book; another patron may need to review two other chapters in the same book and so on. While each individual use is only a small part of the book, the library must digitize the entire book in anticipation of all likely future uses.

207. Authors Guild, Inc. v. HathiTrust, 11 CV 6351 HB, 2012 WL 4808939, at *12 (S.D.N.Y. Oct. 10, 2012).

208. See, e.g., Sony Corp. v. Universal City Studios, Inc., 464 U.S. 419, 449–50 (1984); Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 589 (1994) (remanding “to permit evaluation of the amount taken, in light of . . . considerations of the potential for market substitution”); A&M Records, Inc. v. Napster, Inc., 114 F. Supp. 2d 896, 913 (N.D. Cal. 2000), *aff'd in part, rev'd in part sub nom.* A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001).

209. See *supra* notes 98–144 and accompanying text.

reproduced should be considered much less relevant than in cases where there is a relevant extant market for the work—and reproducing entire works should more likely be considered “neutral” in the fair use analysis—than in cases where a functioning market for the copyrighted work exists.²¹⁰ In *Cambridge University Press v. Becker*, for example, after considering both the amount-and-substantiality factor and the market factor, the court tended to find that Georgia State University could claim as fair use larger amounts of works for which licenses were not reasonably available.²¹¹ The same reasoning applies with force in the case of orphans, for which licenses are decidedly unavailable.

In sum, where libraries and archives are engaging in salutary purposes, and those purposes reasonably require reproducing entire orphaned works, the amount-and-substantiality factor should properly be considered neutral to the fair use analysis.

V. FURTHER STRENGTHENING FAIR USE'S APPLICATION

This Article has shown that nonprofit libraries and archives have strong arguments that digitizing orphan works in their collections and making them available for noncommercial, socially beneficial purposes are fair uses. Of course, the strength of this analysis will vary depending on the facts. A variety of considerations could strengthen fair use's application.

There are at least seven ways libraries and archives could strengthen the fair use arguments for digitizing orphans and making them available to patrons.²¹² Demonstrating good faith²¹³ through record-keeping of reasonable

210. This is broadly in accord with Barton Beebe's empirical findings. Beebe, *supra* note 49, at 615–16.

211. The court individually considered seventy-five book excerpts that were digitized and placed on “electronic reserve.” Where a license to digital versions was not readily available at a reasonable price and in a convenient format, the court found larger percentages to be fair use overall, after also considering the market-harm factor. *Cambridge Univ. Press v. Becker*, 863 F. Supp. 2d 1190, 1232 (N.D. Ga. 2012) (“Factor four weighs in Defendant’s favor when such permissions are not readily available.”). Where a license to digital versions was readily available, however, the court found that ten percent of small or unchaptered books, or one chapter of books with ten chapters or more was “decidedly small,” and justified under the amount-and-substantiality factor; larger portions were not considered “decidedly small” and weighed against fair use. *Id.* at 72–81, 88–89.

212. Some of these suggestions are drawn from case law, and some from practical approaches taken by collections. With regard to the latter, they might best be considered “enhancements” to strengthen a fair use case; I am not arguing that they are necessary. Rather, I am drawing from an idea included in a recent set of practices deemed reasonable by American academic and research libraries. ASSOCIATION OF RESEARCH LIBRARIES, ET AL.,

searches for owners could help establish that owners are indeed unlocatable, and thus unlikely to be harmed by the use. Giving public notice before actually making a digitized orphan available and offering an “opt out” for a period of time could both further demonstrate good faith and more importantly, help avoid inadvertently missing owners who were actually reasonably locatable.²¹⁴ Offering digital copies only of physical works the library holds in its collection, and perhaps limiting the circulation of the physical work while the digital version is available, could help avoid the substitution effects considered under the market-harm factor.²¹⁵ Focusing first on older orphan works in collections, which are both more likely to be orphans and less likely to support a market, might do the same. Offering works only within a library or archive’s patron community, or within a broader research community, could maximize the chance that the use would support salutary purposes under the purpose factor and minimize the chance of economically harming possible owners. For freely available archives, such as online archives that are widely available to the public, limiting the types of uses patrons can make of orphan works to noncommercial, educational, or

CODE OF BEST PRACTICES IN FAIR USE FOR ACADEMIC RESEARCH LIBRARIES (2012), available at http://centerforsocialmedia.org/sites/default/files/documents/pages/code-of-best-practices_arl_0.pdf.

213. “Fair use presupposes ‘good faith’ and ‘fair dealing.’” *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 562 (1985) (citation omitted). As such, courts sometimes explicitly consider the defendant’s good or bad faith in fair use cases. *See, e.g., id.*; *Peter Letterese and Assocs., Inc. v. World Inst. of Scientology Enters.*, 533 F.3d 1287, 1312 n.27 (11th Cir. 2008); *MCA, Inc. v. Wilson*, 677 F.2d 180, 183 (2d Cir. 1981); *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1122 (D. Nev. 2006).

214. The University of Michigan/HathiTrust program has this feature. *Orphan Works FAQs*, UNIV. OF MICHIGAN LIBR., <http://www.lib.umich.edu/orphan-works/faq> (last modified July 24, 2012) (“[A] list of potential orphans will be made available for ninety days before they are opened. If, at any time, a legitimate copyright holder approaches us we will remove their work from our pool of potential orphans.”). From Michigan’s first batch of 160 potential orphans, at least four have been identified as having living authors. *See Orphan Row Update: Another Living Author, Two Books in Print, Literary Estates Held by Charities, Etc.*, THE AUTHORS GUILD (Sept. 15, 2011), <http://blog.authorsguild.org/2011/09/15/orphan-row-update-another-living-author-two-books-in-print-literary-estates-held-by-charities-etc/>. Because the Orphan Works Program was not ripe to be adjudicated, this feature was not directly confronted in the HathiTrust litigation, discussed *supra* notes 17–23. As such, the court did not give guidance as to whether such a system does indeed enhance the legal strength of a fair use claim.

215. The planned University of Michigan/HathiTrust program also had this feature, limiting copies to be made available online to those in the library and further restricting access by keeping “the number of users permitted to view a given work limited at any one time to the number of copies held by” the Michigan library. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *2 (S.D.N.Y. Oct. 10, 2012).

otherwise fair uses²¹⁶ could achieve the same goal. And of course, “taking down” or licensing a work if a licensing market develops and an owner appears could both demonstrate good faith and limit harm to the owner.

VI. CONCLUSION

In sum, fair use has some significant advantages over other approaches through which libraries and archives could make publicly beneficial uses of orphan works. First, a robust fair use defense for such uses of orphans could prevent a substantial societal misallocation of resources. Under fair use, there is no need to pay socially wasteful license fees for works that are not on the market, and for which an owner is unlikely to exist. Further, there is no need for users to obtain *ex ante* permission from a review board or a collecting agency, and no need for the public to invest in developing such a system, significantly reducing administrative and transactional costs.

Second, fair use, as a judicially managed, case-by-case doctrine, has the flexibility to allow for change over time. As libraries and archives discover the best ways to search for owners, preserve works, and make them available to the public, and as new uses develop over time, courts can evaluate whether the changes fit within the parameters of fair use. And perhaps most importantly, allowing fair use of orphans by libraries and archives helps fulfill copyright’s critical purposes of promoting the dissemination of knowledge and supporting speech and free expression.²¹⁷

Relying on fair use has, of course, its own limitations. Fair use is less likely to apply to some uses—such as the commercial sale of orphan works, or the making of commercial derivative works—that can be covered by legislative fixes or collective licenses. In order for a fair use analysis tailored to orphan works to be most applicable, orphans must be appropriately identified, but there are some challenges inherent in this process: searches must be effective in sorting orphans from non-orphans, but cannot be so

216. For an example of such a policy, see the Library of Congress’ description of appropriate uses in its American Memory Collections. *See, e.g., The Wilbur and Orville Wright Papers, supra* note 14:

The Library of Congress provides access to these materials for educational and research purposes and makes no warranty with regard to their use for other purposes. Responsibility for making an independent legal assessment of an item and securing any necessary permissions ultimately rests with persons desiring to use the item.

Id.; *see also* American Women, *supra* note 13.

217. I agree with Professor Samuelson that speech and free expression interests are best considered separate, though related, interests. Samuelson, *supra* note 36, at 2544 n.39.

costly or otherwise onerous as to undermine the benefits of relying on fair use in the first place.

The doctrine's case-by-case nature means that libraries and archives must make risk assessments about whether courts are likely to agree that specific practices constitute fair use, and there is always the possibility that a court would not agree, creating uncertainty. The *Authors Guild v. Hatbi Trust* case clearly offers some recent and useful guidance; at the same time, this case may be appealed, and cannot address all future possibilities. And, of course, finding out with complete certainty whether a given use that has not already been adjudicated as fair is indeed fair would require litigation.²¹⁸ Perhaps most limiting of all is the fact that fair use, in the form analyzed in this

218. For a cogent argument why “complete certainty” is not the appropriate standard for libraries deciding whether or not to rely on fair use, see Kevin L. Smith, *Copyright Risk Management: Principles and Strategies for Large-Scale Digitization Projects in Special Collections*, RESEARCH LIBRARY ISSUES: A QUARTERLY REPORT FROM ARL, CNI, AND SPARC, no. 279 17–23 (June 2012), available at <http://publications.arl.org/rli279/>; see also Kenneth Crews, *If you Cannot Find an Owner*, COPYRIGHT ADVISORY OFFICE, COLUMBIA UNIVERSITY LIBRARY, <http://copyright.columbia.edu/copyright/permissions/if-you-cannot-find-the-owner/> (last visited Sept. 1, 2012). In addition, a small but growing number of scholars have recently challenged the assumption that fair use is unreliable due to its uncertainty, backing their assertions up with data. See, e.g., Beebe, *supra* note 49; Sag, *supra* note 98; Samuelson, *supra* note 36; PETER JASZI & PATRICIA AUFDERHEIDE, RECLAIMING FAIR USE: HOW TO PUT BALANCE BACK IN COPYRIGHT (2011). A number of user communities have recently felt more empowered to rely on fair use in light of Jaszi and Aufderheide's project to create community driven “best practices.” See, e.g., Anthony Falzone & Jennifer Urban, *Demystifying Fair Use: The Gift of the Center for Social Media Statements of Best Practices*, 57 J. COPYRIGHT SOC'Y USA 337 (2010); Michael J. Madison, *Some Optimism about Fair Use and Copyright Law*, 57 J. COPYRIGHT SOC'Y USA 351 (2010); but see Jennifer E. Rothman, *Best Intentions: Reconsidering Best Practices Statements in the Context of Fair Use and Copyright Law*, 57 J. COPYRIGHT SOC'Y USA 371 (2010). The practical availability of the fair use defense cannot be divorced from the question of remedies. How remedies are calibrated affects the risk assessment of a putative fair user—if high remedies create a large downside risk, then uncertainty in the outcome of the defense is more likely to prevent a possible fair user from going forward with a use. Similarly, lower remedies mean that a negative outcome carries with it a smaller downside risk, making it more likely that a possible fair user will move forward in the face of uncertainty. Congress has already protected certain users—“nonprofit educational institution[s], librar[ies], or archives”—from high statutory damages if they have a “reasonable grounds” for believing that a use is fair. 17 U.S.C. § 504(c)(2) (2010). This is an important feature for nonprofit libraries and archives making risk assessments and courts should be willing to apply this feature to all entities engaging in nonprofit library or archival work, not only traditional organizations. I have argued elsewhere that § 504(c)(2)'s protections should be extended beyond these narrow categories of users. Jennifer M. Urban, *Updating Fair Use for Innovators and Creators in the Digital Age: Two Targeted Reforms* (Public Knowledge Fair Use Whitepaper, Feb. 15, 2010), available at <http://www.publicknowledge.org/fair-use-whitepaper>; see also Katz, *supra* note 106.

Article, is a nearly unique feature of U.S. copyright law and does not apply in other countries.²¹⁹

As such, fair use cannot cover the entire landscape of uses people might make of orphan works, and other solutions under discussion might better address a variety of uses. But for the digitization projects of nonprofit libraries and archives in the United States, its benefits make fair use a powerful tool for freeing the orphans in their collections.

219. In recent years, however, some other countries have begun adding significant flexibility into their copyright regimes, including some fair-use-like provisions. *See, e.g.*, Copyright Act, 5768-2007, 2007 LSI 34, 19 (2007) (Isr.), *available at* http://www.wipo.int/wipolex/en/text.jsp?file_id=132095; Intellectual Property Code of the Philippines § 185, Rep. Act 8293 (Phil.), *available at* http://www.wipo.int/wipolex/en/text.jsp?file_id=129343; Copyright Act, part III, div. 3 (2006) (Sing.), *available at* http://www.wipo.int/wipolex/en/text.jsp?file_id=187736. Israel's provision, in particular, is very similar to United-States-style fair use. It follows § 107 exactly, except that it eliminates the reference to "whether such use is of a commercial nature or is for nonprofit educational purposes" under the first factor. Copyright Act, 5768-2007, 2007 LSI 34, 19 (2007–2008) (Isr.), *available at* http://www.wipo.int/wipolex/en/text.jsp?file_id=132095. The Philippines law is also very similar. Intellectual Property Code of the Philippines § 185, Rep. Act 8293 (Phil.), *available at* http://www.wipo.int/wipolex/en/text.jsp?file_id=129343.

ABANDONING THE ORPHANS: AN OPEN ACCESS APPROACH TO HOSTAGE WORKS

Lydia Pallas Loren[†]

ABSTRACT

The lengthening of the duration of copyright protection and the elimination of copyright registration formalities have contributed to the rise of so-called orphan works: works that remain subject to copyright law but whose owners cannot be identified and located by someone who wishes to make use of the work in a manner that requires permission of the copyright owner. In this Article I examine one of the root causes of the inability to address the orphan work problem: the metaphor of the “orphan” itself. I propose that these works should not be viewed as orphans, but rather as “hostages”—constrained in their movement by the restricting combination of the set of rules established by copyright law and the absence of the owner who could release the works from what binds them in their confinement.

The hostage metaphor leads to a clearer recognition that what is needed is not a stand in for the “parent” of these orphans, rather what is called for is an incentive for responsible parties to operate as “special forces” to free the hostages. I propose a limited immunity for entities that act as “special forces” in freeing the hostages. The immunity should be available when an entity non-negligently identifies a work as a hostage work and provides an open access copy of that work with the hostage-freeing information attached in human and machine readable form. I also suggest that courts should employ equitable doctrines to limit the infringement remedies available against derivative work creators that would discourage copyright owners from delaying in making their presence and availability known. Adopting this approach would provide an appropriate level of protection for copyright owners and significantly reduce a form of waste created by the hostage work problem.

© 2012 Lydia Pallas Loren.

[†] Professor of Law and Kay Kitagawa and Andy Johnson-Laird Faculty Scholar, Lewis & Clark Law School. I thank Tricia Juettemeyer for her research assistance, participants at the 2012 Works-In-Progress in Intellectual Property (WIPIP) workshop and at a 2012 Lewis & Clark Faculty Workshop for helpful feedback on earlier iterations of this project, Tomás Gómez-Arostegui for insightful suggestions on a earlier draft of this Article, and Pamela Samuelson for her invitation to participate in this symposium and her ongoing encouragement on this project.

TABLE OF CONTENTS

I.	INTRODUCTION.....	1432
II.	EXAMINING THE ORPHAN METAPHOR AND ALTERNATIVES	1435
	A. ORPHANS.....	1435
	B. ABANDONED AND NEGLECTED PROPERTY	1439
	C. ADVERSE POSSESSION	1443
	D. PRESCRIPTIVE EASEMENTS.....	1448
	E. DOCTRINE OF WASTE.....	1449
	F. HOSTAGE WORKS.....	1451
III.	FREEING HOSTAGE WORKS	1455
	A. PROTECTING RESPONSIBLE “SPECIAL FORCES”	1456
	1. <i>Responsible Identification of Hostage Freeing Information</i>	1459
	2. <i>Requiring Open-Access to Works with Accompanying Freeing Information</i>	1462
	B. HOSTAGE WORKS AND DERIVATIVE WORK CREATORS.....	1466
	1. <i>Injunctive Relief</i>	1467
	2. <i>Statute of Limitations</i>	1468
	3. <i>Laches</i>	1469
IV.	CONCLUSION	1470

I. INTRODUCTION

Many aspects of the copyright protection systems in the United States and other countries have contributed to the troubling situation now known as the “orphan work problem.” The lengthening duration of copyright protection—along with the elimination of the requirements for copyright registration, copyright notices on published copies, and renewal registrations—has contributed to the rise of so-called “orphan works”:¹ works that remain subject to copyright law but whose owners “cannot be identified and located by someone who wishes to make use of the work in a manner that requires permission of the copyright owner.”²

1. David R. Hansen, *Orphan Works: Causes of the Problem* (Berkeley Digital Library Copyright Project, White Paper No. 3, 2012), available at <http://ssrn.com/abstract=2038068>.

2. REGISTER OF COPYRIGHTS, U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 1 (2006), available at <http://www.copyright.gov/orphan/orphan-report.pdf> [hereinafter REPORT ON ORPHAN WORKS]; see also Orphan Works Notice of Inquiry, 70 Fed. Reg. 3,739, 3,739 (Jan. 26, 2005) (describing orphan works as those “whose owners are difficult or even impossible to locate”).

Orphan works result from market failure in which a potential user of a copyrighted work faces insurmountable transaction costs to obtain authorization from the rights holder. Facing the risk of infringement liability, the user foregoes the use, even though the parties would have struck a deal for that use had the user been able to locate the copyright owner.

Proposals abound for dealing with the orphan work problem.³ In 2006, the U.S. Copyright Office issued a massive report.⁴ In 2008, proposed legislation passed the U.S. Senate,⁵ but failed to win passage in the House of Representatives.⁶ In 2011, the European Commission proposed a directive addressing the orphan work problem.⁷ Today the class-action styled litigation over the Google Book Search project, the proposed settlement of which would have addressed aspects of the orphan work problem, remains in limbo.⁸ A recent lawsuit filed against the HathiTrust, a consortium of some of the country's leading research institutions and libraries, which sought to provide access to works it identified as "orphans," is in the earliest phases of litigation.⁹ And scholars continue to propose reforms to eliminate or at least ameliorate the problem.¹⁰

A variety of competing dynamics contribute to the inability to address the orphan work problem. The stakeholders in a solution include the wide array of potential "users"¹¹ of copyrighted works: libraries, museums,

3. For an overview of the variety of proposed solutions, see David R. Hansen, *Orphan Works: Mapping the Possible Solution Spaces* (Berkeley Digital Library Copyright Project, White Paper No. 2, 2012), available at <http://ssrn.com/abstract=2019121>.

4. See REPORT ON ORPHAN WORKS, *supra* note 2.

5. S. 2913, 110th Cong. (2008).

6. Orphan Works Act of 2008, H.R. 5889, 110th Cong. (2008).

7. *Proposal for a Directive of the European Parliament and of the Council on Certain Permitted Uses of Orphan Works*, at 10, COM (2011) 289 final (May 24, 2011), available at http://ec.europa.eu/internal_market/copyright/orphan_works_en.htm [hereinafter *Proposal for a Directive*]. In October 2012, as this Article was preparing to go to press, the European Union approved a final directive. Directive 2012/28, of the European Parliament and of the Council on Certain Permitted Uses of Orphan Works, 2012 O.J. (L 299) [hereinafter *Final Directive*].

8. See *Authors Guild v. Google Inc.*, 770 F. Supp. 2d 666, 669 (S.D.N.Y. 2011) (rejecting the Amended Settlement Agreement).

9. The lawsuit, *Authors Guild, Inc. v. HathiTrust*, No. 11-CIV-6351 (S.D.N.Y. filed Sep. 12, 2011), was filed in 2011. In October, the district court granted the HathiTrust's motion for judgment on the pleadings concerning its Orphan Works Project because of concerns over the claims' ripeness for adjudication. See *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *8 (S.D.N.Y. Oct. 10, 2012).

10. Indeed, in this Article I provide my own suggestions for reforms.

11. The terminology of "users" in the copyright context has its own metaphor power. The other instance in which the term "users" is used routinely is to describe those who are addicted to illegal drugs.

archives, educational institutions, movie producers, educators, and anyone else who would like to build on the expressive works of the past, as well as owners of copyrights in works that might erroneously be labeled as orphans. As James Grimmelmann noted, in the ongoing orphan work debate, “hope, fear, and anger will tug at stakeholders in subtle and complicated ways.”¹²

In this Article I examine one obstacle to addressing the orphan work problem: the metaphor of the “orphan” itself.¹³ I begin in Part II with an exploration of the orphan metaphor and potential alternative metaphors that could be used to identify these works. A variety of doctrines in real and chattel property provide interesting insights that help inform the approach proposed later in the Article. I conclude that these works should not be viewed as “orphans,” but rather as “hostages”—constrained in their movement by the restricting combination of copyright law and the absence of copyright owners who could release works from their confinement.

The hostage metaphor leads to a clearer recognition that stand-ins for the “parents” of these “orphans” are not the solution. Rather this problem calls for an incentive to encourage responsible parties to operate as “special forces” to free the “hostages.” Part III proposes limited immunity for entities that act as “special forces” in freeing the hostages by, first, responsibly generating and disseminating reliable information concerning the copyright status of a work, and second, by providing access to a digital copy of the work that adheres to basic open access principles. Part III also urges courts to employ equitable doctrines to limit the remedies available against derivative work creators so as to discourage copyright owners from delaying in making their presence and availability known.

As further developed in Part III, I offer this open access model in a way that could be implemented either legislatively or judicially, through a combination of the ever-malleable fair use balancing test and a variety of equitable doctrines. This model would affect the remedies that courts grant in lawsuits brought by resurfacing copyright owners. The system has the potential to facilitate the development and dissemination of reliable information, increase the public benefit through greater access to hostage works, and encourage calculated risk-taking by derivative work makers. Such a system would provide appropriate protection for copyright owners and at the same time significantly reduce a form of waste created by the hostage work problem.

12. James Grimmelmann, *The Orphan Wars*, *EDUCAUSE REV.*, Jan. 23, 2012, <http://www.educause.edu/ero/article/orphan-wars>.

13. See WILLIAM PATRY, *MORAL PANICS AND THE COPYRIGHT WARS* 76–78 (2009).

II. EXAMINING THE ORPHAN METAPHOR AND ALTERNATIVES

Metaphors can have powerful effects in both identifying problems and in shaping approaches to address them. Indeed, as others have observed, “metaphors are not merely literary, but cognitive. Language maps deep cognitive structures and thus mirrors cultural patterns and social structures.”¹⁴ The best persuaders, including lawyers and lobbyists, have long exploited the strategic use of metaphors. Examining the orphan metaphor exposes the important ways in which it has swayed our opinions both of what the problems are and what the solutions might be. Exploring alternative metaphors can productively expand thinking on this critical problem in copyright law.

A. ORPHANS

The choice to call certain works “orphan” works fits comfortably within another significant and longstanding image in copyright: the romantic author. “The romantic author is the individuated figure for whose benefit, reward, and encouragement the conferral and expansion of IP rights has been justified.”¹⁵ The vision of the romantic author creating works of genius that deserve protection has been a common story used to support the notion of granting copyright in the first place.¹⁶

In the metaphor of the romantic author, the works he creates are his children, born of his labor and genius.¹⁷ He receives the right to control, and thus protect, his children. Using the word “orphans” to describe works whose copyright owners cannot be located pulls on that metaphor and triggers the concerns any humane person would have toward abandoned children. These orphans have suffered the tragic loss of their parents. These are works whose parents have been lost or killed, or whose parents have long

14. George H. Taylor & Michael J. Madison, *Metaphor, Objects, and Commodities*, 54 CLEV. ST. L. REV. 141, 159 (2006).

15. Keith Aoki, *Distributive and Syncretic Motives in Intellectual Property Law (with Special Reference to Coercion, Agency, and Development)*, 40 U.C. DAVIS L. REV. 717, 732 (2007).

16. See Anupam Chander & Madhavi Sunder, *The Romance of the Public Domain*, 92 CALIF. L. REV. 1331, 1339 (2004); Mark A. Lemley, *Romantic Authorship and the Rhetoric of Property*, 75 TEX. L. REV. 873, 886 (1997) (reviewing JAMES BOYLE, SHAMANS, SOFTWARE, AND SPLEENS: LAW AND CONSTRUCTION OF THE INFORMATION SOCIETY (1997)).

17. In the romantic visions, the author is always male. See, e.g., Rosemary J. Coombe, *The Properties of Culture and the Politics of Possessing Identity: Native Claims in the Cultural Appropriation Controversy*, 6 CAN. J.L. & JURIS. 249, 250 n.7 (1993); Carys J. Craig, *Reconstructing the Author-Self: Some Feminist Lessons For Copyright Law*, 15 AM. U. J. GENDER SOC. POL'Y & L. 207, 242 (2007).

ago abandoned them.¹⁸ We reflexively begin to believe that orphan works need the kind of protection that society provides to abandoned children.¹⁹

The state protects orphaned children for a certain amount of time, until they reach the age of maturity or are able to care for themselves.²⁰ The state stands *in loco parentis*, tasked with protecting the orphans from the evils of the world. The perceived evils of the world are multifold and include Dickensian images: bleak orphanages, barren workhouses, and street gangs assembled by Artful Dodger of *Oliver Twist*, where the unfortunate children are put to work for commercial entities exploiting whatever commercial value can be obtained from the children's labor.²¹ These orphan exploiters fail to invest in or care for the children properly, and yet usurp the work-value of the orphan child. This implied narrative of the potential abuse has impeded the passage of orphan works legislation.

Orphan works, however, are quite different from abandoned children.²² First, the "parent" of the orphan work may not, in fact, be a single romantic author. Copyrighted works often are created through collaboration and often are prepared in an employment context.²³ In many instances the "parent," or author, is not an individual at all, but rather a corporation or other non-

18. PATRY, *supra* note 13, at 77.

19. *Id.*; see Taylor & Madison, *supra* note 14, at 162 (noting: "[i]f copyrighted works implicitly have 'parents,' rather than authors and publishers, the 'orphan' label is consistent with viewing the orphan works problem much as child welfare advocates view the problem of abandoned and abused children"). That desire to "protect" the "orphans" may be part of what has led to proposals for regimes that require licensing fees for the use of orphan works be placed into a kind of escrow, see Hansen *supra* note 3, at 16–18, and to calls for the court to appoint representatives to advocate on behalf of the absentee "parents," see, e.g., Amended Settlement Agreement at 82, Authors Guild, Inc. v. Google Inc., No. 05 CV 8136 (S.D.N.Y. Nov. 13, 2009), available at <http://www.googlebooksettlement.com/intl/en/AmendedSettlement-Agreement.zip> (defining "Unclaimed Works Fiduciary").

20. See, e.g., Or. Rev. Stat. §§ 418.005, 418.015, 418.027 (2012). All fifty U.S. states have systems in place to protect orphaned children. See CHILD WELFARE INFORMATION GATEWAY, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, ONLINE RESOURCES FOR STATE CHILD WELFARE LAW AND POLICY (2011), available at http://www.childwelfare.gov/systemwide/laws_policies/statutes/resources.pdf.

21. CHARLES DICKENS, THE ADVENTURES OF OLIVER TWIST 5, 11–12, 57–60, 413 (Boston: Ticknor and Fields 1866) (1838), available at <http://books.google.com/books?id=1bMXAAAAAYAAJ>.

22. Lior Jacob Strahilevitz, *The Right to Abandon*, 158 U. PA. L. REV. 355, 380–84 (2010) (discussing the intentional abandonment of children).

23. Catherine Fisk, *Authors at Work: The Origins of the Work-for-Hire Doctrine*, 15 YALE J.L. & HUMAN. 1 (2003).

human entity.²⁴ If works are not created by romantic solo authors, their creations are not “children” that may be orphaned upon the author’s demise.

Further, what matters in the context of obtaining permission for a particular use is not authorization from the author or “parent” of a work. Rather, when someone seeks to use a work in a manner that would constitute infringement, authorization of the copyright *owner* is required; authorization from the “parent” author is irrelevant unless the author has retained ownership of the copyright.²⁵ While in the United States copyright ownership vests initially in the author,²⁶ the Copyright Act expressly provides for the ability to transfer the copyright rights in whole or in part.²⁷ Some works become “orphans” through the death of the individual author who created the work and retained ownership of the copyright. Other works, however, are “orphaned” because the copyright had long ago been sold and either the assignee entity no longer exists or the chain of title can no longer be traced.²⁸ The important legal distinctions between the individuals who create a work, the legal author of a work, and the copyright owner of a work, are lost when the orphan metaphor is employed.²⁹

Orphan works are also quite different from abandoned or orphaned children with respect to what occurs upon maturation. When an orphaned child reaches the age of maturity, he or she assumes control of, and becomes responsible for, his or her own caretaking and destiny. The maturation of a copyrighted work, on the other hand, results in the work entering the public domain and its entire expressive content becoming part of the storehouse of knowledge.³⁰ A “matured” orphan work is thus freely available for others to

24. In the United States, the authors of works prepared by employees in the scope of their employment are not the employees but rather the employers—whether they are corporations, partnerships, or other entities. *See* 17 U.S.C. § 201(b) (2010) (specifying that in the case of “works made for hire,” the “author” is the hiring party); *see also id.* § 101 (defining a “work made for hire” to include works made by an employee within the scope of employment).

25. *See, e.g., American Geophysical Union v. Texaco*, 802 F. Supp. 1, 27 (S.D.N.Y. 1992) (stating that “[o]nce an author has assigned her copyright, her approval or disapproval of [defendant’s use] is of no further relevance”).

26. 17 U.S.C. § 201(a).

27. *Id.* § 201(d).

28. REPORT ON ORPHAN WORKS, *supra* note 2, at 26–29. The divisibility of the rights granted to a copyright owner further exacerbates the orphan work problem.

29. Even the Copyright Office itself uses the paired parent/orphan metaphor without careful attention to this distinction. *See, e.g., id.* at 34 (stating that the term orphan “certainly must mean what it implies: that the ‘parent’ of the work is unknown or unavailable”).

30. Even prior to the expiration of copyright protection, the ideas and other unprotected elements of the copyrighted work are already part of copyright’s public domain. 17 U.S.C. § 102(b). “[E]very idea, theory, and fact in a copyrighted work becomes instantly

copy and build upon. An orphan work's passage into the public domain permits users to incorporate expression from the older work into new expressive works of authorship. In this way, a matured copyrighted work is "owned" by the public and its use cannot be constrained.³¹

While the orphan metaphor may be misleading, the metaphor does not appear to have been in use for all that long.³² The term first gained traction in the mid-1990s to describe neglected film footage that was at serious risk of loss due to degradation in the physical medium over time.³³ In the film preservation context, the term "orphan" applied to films whose copyright owners could not be located, but it also encompassed films with other characteristics.³⁴ The first official use of the phrase to identify copyrighted

available for public exploitation at the moment of publication." *Eldred v. Ashcroft*, 537 U.S. 186, 219 (2003). See also Edward Samuels, *The Public Domain in Copyright Law*, 41 J. COPYRIGHT SOC'Y 137, 137 (1993) (suggesting that public domain is what remains when all forms of protected information are taken into account); Pamela Samuelson, *Mapping the Digital Public Domain: Threats and Opportunities*, 66 LAW & CONTEMP. PROBS. 147, 148–51 (2003).

31. The Supreme Court recently recognized the constitutionality of Congress' ability to provide exclusive rights to works that previously had entered the public domain. *Golan v. Holder*, 132 S. Ct. 873, 894 (2012). An act of Congress is, however, required, and circumstances under which this maneuver is constitutional appear at least somewhat constrained. *Id.* at 887 (discussing the limited reasons for congressional action of this type).

32. In the Westlaw Journals and Law Reviews database, the first published appearance of the phrase "orphan work" is in 2000—in two places. First, in Jane C. Ginsburg et al., *The Constitutionality Of Copyright Term Extension: How Long Is Too Long?*, 18 CARDOZO ARTS & ENT. L.J. 651, 666 (2000), the phrase appears in the context of the full reproduction of the appellant's reply brief in *Eldred v. Reno*, 74 F. Supp. 2d 1 (D.D.C. 1999). Also published in 2000 was Marybeth Peters, *Copyright Office Update*, 599 PLI/PAT 185, 242 (2000), in which Ms. Peters, then the Register of Copyrights for the United States, indicates a problem identified by educators who are unable to locate copyright owners to negotiate with, and stating that the problem of "orphan works" may become acute due to the longer term of copyright. This is in the "Licensing Issues" section of her *Update*. *Id.* at 241. While Ms. Peters provides no citation for the reference to "orphan works," the phrase appears in a report issued by the Copyright Office the previous year. See U.S. COPYRIGHT OFFICE, REPORT ON COPYRIGHT AND DIGITAL DISTANCE EDUCATION, xxiv (1999), available at http://www.copyright.gov/reports/de_rprt.pdf.

33. LIBRARIAN OF CONGRESS, FILM PRESERVATION 1993: A STUDY OF THE CURRENT STATE OF AMERICAN FILM PRESERVATION, available at <http://www.loc.gov/film/study.html>.

34. The 1993 report on film preservation issued by the Librarian of Congress used the term "orphan" in describing the category of films that were in need of preservation:

If there is a single division that separates most of the preservation issues discussed in this report, it is between two categories of films: those that have evident market value and owners able to exploit that value; and the other films, often labeled "orphans," that lack either clear copyright holders or commercial potential to pay for their continued preservation. In practice, the former are primarily features from major Hollywood studios; the latter—numerically the majority—include newsreels and

works whose copyright owners were missing did not appear until 1999.³⁵ Thus, while the metaphor enjoys widespread use today, the short duration of its use may permit a shift in terminology if a better metaphor were available. The next Sections explore candidates for alternative metaphors that could be employed to better understand the “orphan works” problem.

B. ABANDONED AND NEGLECTED PROPERTY

Without the “orphan” metaphor, what are these works whose owners cannot be located? One possible alternative metaphor is that these works are abandoned, neglected, or derelict property.³⁶ As abandoned or neglected property, these works have value that is being wasted by the absentee owner. The value is apparent because, by definition of the problem, there is a user desiring to engage in a use of the work but the user cannot locate the owner to obtain authorization. In the area of tangible property, the common law developed a variety of doctrines designed to minimize the waste that results from abandoned or neglected property. Specifically, a variety of legal rules established through common law and legislation seek to transition ownership in situations where property is not being cared for.³⁷ The new owners facilitate market transactions concerning the property. Examining these doctrines and statutory enactments provides insight into how an alternative

documentaries, avant-garde and independent productions, silent films where copyright has expired, even certain Hollywood sound films from now defunct studios.

Id. In a 1999 symposium hosted by the University of South Carolina, titled: “Orphans of the Storm: Saving ‘Orphan Films’ in the Digital Age,” <http://www.sc.edu/filmsymposium/archive/orphans1999/details.html> (last visited June 30, 2012), Paolo Cherchi-Usai asserted that the first use of the term was in 1993 by David Francis, then Chief of the Motion Picture, Broadcasting, and Recorded Sound Division at the Library of Congress, at the Los Angeles hearing for the National Film Preservation Plan. *See* Paolo Cherchi-Usai, What is an Orphan Film? Definition, Rationale and Controversy (Sept. 23, 1999) (transcript available at <http://www.sc.edu/filmsymposium/archive/orphans1999/2.1%20Paolo%20Cherchi-Usai.htm>); *see also* Press Release, San Francisco Film Society, David Francis to Receive Mel Novikoff Award at the 45th San Francisco International [sic] Film Festival (2002), *available at* http://www.sfiff.org/press/novikoff_02.html. While the term “orphans” had been used in the context of copyright law previously, it was to discuss works that were not eligible for copyright protection. *See, e.g.*, Leslie J. Hagin, *A Comparative Analysis of Copyright Laws Applied to Fashion Works: Renewing the Proposal for Folding Fashion Works into the United States Copyright Regime*, 26 TEX. INT’L L.J. 341 (1991); Natalie Wargo, *Copyright Protection for Architecture and the Berne Convention*, 65 N.Y.U. L. REV. 403, 452 (1990).

35. *See supra* note 33.

36. PATRY, *supra* note 13, at 76–77.

37. The taxation of real property can be understood in this light. When an owner of property fails to pay property taxes, the government has the ability to seize the property and sell it to an owner that will pay those taxes. Strahilevitz, *supra* note 22, at 394–98.

metaphor might assist us in thinking differently about the problem of copyrighted works whose owners cannot be located, and about the appropriate solutions.

In examining doctrines related to tangible property rights to illuminate thinking about designing doctrines for intangible rights, the fundamental differences between these two types of rights must be kept in mind. Among other ends, legal recognition of tangible property rights seeks to avoid over-consumption of, and under-investment in, property (a.k.a. the tragedy of the commons).³⁸ The intangible copyrighted work, on the other hand, does not face the same problems of potential over-consumption. Copyrights and other intellectual property assets lack inherent characteristics of excludability and rivalrous consumption. In fact, these assets are characterized by non-rivalrous consumption.³⁹ For these types of “property” rights, the risk of over-consumption does not animate the creation of property rights in the first place.⁴⁰ Rather, the utilitarian justification for copyright protection is one based on a fear of under-investment in the creation and dissemination of the expression.⁴¹ Copyrights are granted in an effort to ensure the incentive to invest in the creation and dissemination of expressive works.⁴² Mindful of these important differences in the fundamental characteristics of intangible and tangible property, we can better evaluate whether it would be more appropriate to call this problem the “abandoned works problem.”

Abandoned chattel property can present problems for society. Abandoned chattel property with little or negative value, a.k.a. garbage, will unlikely be claimed by a new owner. The cost of abandoned garbage includes the health consequences if not removed promptly and properly, along with hauling and disposal costs. To minimize these problems, governments adopt laws that prohibit abandonment of negative value property, such as anti-littering laws. Abandoned chattel property with positive value can result in what one scholar has termed “lawless-race” costs, as would-be claimants (a.k.a. “finders”) seek to gain possession, and ownership, of the abandoned

38. Garrett Hardin, *The Tragedy of the Commons*, 162 *SCIENCE* 1243 (Dec. 13, 1968) (introducing the term).

39. Julie Cohen, et al., *COPYRIGHT IN A GLOBAL INFORMATION ECONOMY* 6–7 (3d Ed. 2010).

40. Lydia Pallas Loren & Joseph Scott Miller, *INTELLECTUAL PROPERTY LAW: CASES AND MATERIALS* 2–3 (Ver. 2.2 2011).

41. *Id.*; see also WILLIAM M. LANDES & RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW* 11, 13–14 (2003).

42. *Id.* at 56–57.

property.⁴³ To reduce the possibility of suffering the cost of the “lawless race,” laws establish rules governing title to the property.

One way to handle the potential negative effects of abandoned positive-value chattel is for title to escheat to the state, permitting the state to control the transfer of title to a new owner. Both the Uniform Unclaimed Property Act of 1995⁴⁴ and the Federal Abandoned Shipwreck Act of 1987⁴⁵ employ this approach. Interestingly one justification for the escheat rule in the case of abandoned shipwrecks is the belief that the state, rather than the first finder, will have a stronger incentive to preserve the shipwreck for research and historic purposes, as opposed to plundering the shipwreck’s assets and selling them on the open market.⁴⁶

What would happen if we referred to copyrighted works whose owners cannot be located as “abandoned works”? Calling such works “abandoned works” might lead us to desire solutions designed to identify a new owner—either we would entrust the work to the state to protect its cultural or historic value,⁴⁷ or give a new private entity the opportunity to put the work to good use. A likely candidate for the new owner of an “abandoned work” would be the party who seeks to use the copyrighted work but who cannot locate the owner. Concerned with lawless-race costs, this solution might make some sense. However, with “abandoned” copyrighted works of positive value, we do not face the costs of potential lawless races because of the non-rivalrous consumption characteristic. More than one person or entity can make use of a copyrighted work without interfering with another’s use.⁴⁸

In the context of chattel property, the distinction between negative value goods, *i.e.* trash, versus positive value goods, *i.e.* treasure, helps identify the costs of abandoned property: the health and disposal costs in the context of

43. Strahilevitz, *supra* note 22, at 409–11.

44. As of June 30, 2012, fifteen states and the U.S. Virgin Islands have adopted the Uniform Unclaimed Property Act of 1995. See Uniform Law Commission, *Legislative Fact Sheet—Unclaimed Property Act*, <http://uniformlaws.org/LegislativeFactSheet.aspx?title=Unclaimed%20Property%20Act> (last visited June 30, 2012).

45. 43 U.S.C. §§ 2101–2106 (2010).

46. See Shelly R. McGill, *Are Criticisms of the Abandoned Shipwreck Act Anchored in Reality?*, 29 ENVIRONS ENVTL. L. & POL’Y J. 105, 117 (2006).

47. Interestingly, this is consistent with the context in which the metaphor of “orphan works” was first employed—to support the creation of the National Film Preservation Board whose “primary mission is to save orphan films . . . the living record of the twentieth century.” See National Film Preservation Board, *About the Board*, <http://www.loc.gov/film/filmabou.html> (last updated Aug. 31, 2011).

48. In this context it is important to not confuse tangible *copies* of a copyrighted work with the intangible work. The copies have rivalrous consumption characteristics common to all tangible property.

negative value goods and the costs of lawless races and waste in the context of positive value goods. Should we characterize copyrighted works for which the copyright owners cannot be located as negative value or positive value abandoned works? Interestingly, part of the problem with calling these works “orphans” is that the metaphor obscures whether each individual orphan work is trash or treasure.⁴⁹ Consequently, we have difficulty recognizing and evaluating the costs that society suffers. However, the “problem” of “orphan works” arises when there exists a user who desires to make a use of a copyrighted work. The fact that there is a new user desirous of making a use indicates that the work retains at least some positive value.

With regard to “abandoned” copyrighted works, while we do not face the costs associated with lawless races, we do face a problem of waste—these works are not being disseminated but they should be. Society endures the costs associated with waste without any offsetting benefit gained by continuing to respect the copyright rights of the absent owner. If someone could grant permission for the use, then the waste could be avoided or at least ameliorated. Rules that establish a new owner facilitate market-based transactions. An individual desiring to engage in a use that requires permission would be able to determine who could grant that permission. Identifying a new owner does not further copyright’s incentive for creation, although it might facilitate an incentive for subsequent dissemination.⁵⁰ Adverse possession of real property, discussed below, has a similar dynamic.

The fact that copyright law already has an abandonment doctrine presents a major hurdle to calling works whose copyright owner cannot be located “abandoned works.” Abandonment extinguishes the copyright interest.⁵¹ To abandon a copyright in a work the owner must intend to surrender the rights and must commit some overt act showing this intent.⁵² Courts have been quite stringent in applying these requirements: a mere loss of interest in exploiting the exclusive rights is not sufficient to give rise to

49. “The Congressional Research Service has estimated that just two percent of copyrighted works that are 55 to 75 years old retain any commercial value.” Lawrence Lessig, *Little Orphan Artworks*, N.Y. TIMES, May 20, 2008, at A23.

50. Golan v. Holder, 132 S. Ct. 873, 888 (2012) (discussing the importance of an incentive for dissemination in addition to an incentive for creation).

51. See Lydia Pallas Loren, *Building a Reliable Semicommons of Creative Works: Enforcement of Creative Commons Licenses and Limited Abandonment of Copyright*, 14 GEO. MASON L. REV. 271, 319–22 (2007).

52. Nat’l Comics Publ’ns, Inc. v. Fawcett, Inc., 191 F.2d 594, 598 (2d Cir. 1951); Hadady Corp. v. Dean Witter Reynolds, Inc., 739 F. Supp. 1392, 1395–96, 1398–99 (C.D. Cal. 1990).

abandonment.⁵³ Given the strict requirements of the abandonment doctrine in copyright law and the consequences of a determination of abandonment, courts have deemed very few works to have been abandoned.⁵⁴

When an owner abandons the copyright in a work, that work enters the public domain. This result comes from the notion that the intangible rights protected by copyright, while referred to as intellectual *property* rights, are merely exclusive rights Congress grants for limited times. If an owner no longer is interested in those exclusive rights, it does not follow that another should be granted rights. Instead, the work should enter the public domain faster than it would in due course.⁵⁵ If works whose copyright owners cannot be found were referred to as “abandoned works,” the solution to the “problem” would seem clear: courts should deem these works to be in the public domain.⁵⁶

C. ADVERSE POSSESSION

The doctrine of adverse possession merits exploration as a doctrine relevant when an owner of real property⁵⁷ is neglectful, absent, or difficult to locate. Adverse possession provides a possessor of property with a mechanism to obtain legal title to the property if the possessor can demonstrate the required elements: possession must be (1) actual, (2) open

53. See generally Matthew W. Turetzky, *Applying Copyright Abandonment in the Digital Age*, 19 DUKE L. & TECH. REV. ¶¶ 21–24 (2010) (exploring the doctrine of copyright abandonment and reviewing cases).

54. *Id.* ¶¶ 22–23.

55. Trademark law is similar: “When a trademark falls into disuse, there is no longer any justification for impoverishing the public domain . . . so the mark is returned to the commons where it can be appropriated by any other firm that wishes to use it in commerce.” Strahilevitz, *supra* note 22, at 391.

56. Courts’ resistance to declaring works to be in the public domain results in part from the nature of litigation: the copyright owner is asserting her rights by litigating an infringement action, yet in that action the defendant is seeking a declaration that the copyright owner has abandoned her rights. It is difficult to rule a copyright owner has abandoned her copyright when she is litigating a case of alleged infringement.

57. Adversely possessing chattel property generally follows the same requirements of adverse possession of real property. However, a fundamental difference is that an adverse possessor’s use of personal property can be open and notorious and yet even a diligent owner may not be able to obtain notice of the adverse claim. Patty Gerstenblith, *The Adverse Possession of Personal Property*, 37 BUFF. L. REV. 119, 124 (1989). This difficulty arises from the fact that the statute of limitations for asserting one’s claim begins to run only once the owner knows or should know the location and possessor of the property. See *O’Keeffe v. Snyder*, 416 A.2d 862 (N.J. 1980); *Solomon R. Guggenheim Found. v. Lubell*, 569 N.E.2d 426 (N.Y. 1991) (holding that a statute of limitations does not begin to run until a demand for return of the property is made to a good faith purchaser). Therefore, adverse possession of chattel property is difficult to accomplish.

and notorious, (3) exclusive, (4) continuous, (5) hostile under a claim of right, and such possession must have occurred for a time that exceeds the statute of limitations for an ejection action.⁵⁸

One way to understand the doctrine of adverse possession is that it seeks to penalize the neglectful owner. Sometimes referred to as the “demerit” theory of adverse possession,⁵⁹ this view of the doctrine penalizes the negligent or dormant owner for “sleeping upon his rights.”⁶⁰ By creating a risk of loss of title, the doctrine establishes an incentive for owners of real property to remain at least minimally attentive to their assets.

Another way to understand the doctrine is that it seeks to reward the industrious possessor. Sometimes referred to as the “merit” theory of adverse possession, this view of the doctrine creates an incentive for possessors to productively use the property.⁶¹ Because the use must be open, notorious, and continuous for the relevant time period, only one who is using the property will be able to obtain legal title. In an important study of judicial application of the doctrine, it appears courts may also require, implicitly if not explicitly, the subjective good faith of the adverse possessor.⁶² Requiring good faith, or at least the absence of bad faith, is in keeping with a “merit” theory of adverse possession.

A more prevalent and accepted view of the doctrine, offered by Professor Ballantine almost a century ago, is that adverse possession “has not for its object to reward the diligent trespasser for his wrong nor yet to penalize the negligent and dormant owner for sleeping upon his rights; the great purpose is automatically to quiet all titles which are openly and consistently asserted, to provide proof of meritorious titles, and correct errors in conveyancing.”⁶³ In other words, adverse possession helps to identify with whom someone can deal if they desire to make use of property.⁶⁴ If an adverse possessor’s use is open for all to see and the prior

58. See Thomas W. Merrill, *Property Rules, Liability Rules, and Adverse Possession*, 79 NW. U. L. REV. 1122, 1123 (1985).

59. See Henry W. Ballantine, *Title by Adverse Possession*, 32 HARV. L. REV. 135, 135 (1918) (quoting JAMES BARR AMES, LECTURES ON LEGAL HISTORY: AND MISCELLANEOUS LEGAL ESSAYS 197 (1913)) (internal quotation marks omitted) (“English lawyers regard not the merit of the possessor, but the demerit of the one out of possession.”); see also Matthew W. Daus, *The Adverse Possession of Copyright*, 13 LOY. L.A. ENT. L. J. 45, 86 (1992).

60. Ballantine, *supra* note 59, at 135 (quoting AMES, *supra* note 59, at 197).

61. John G. Sprankling, *An Environmental Critique of Adverse Possession*, 79 CORNELL L. REV. 816 (1994).

62. R.H. Helmholz, *Adverse Possession and Subjective Intent*, 61 WASH. U.L.Q. 331, 347 (1983).

63. Ballantine, *supra* note 59, at 135.

64. Carol M. Rose, *Possession as the Origin of Property*, 52 U. CHI. L. REV. 73, 78 (1985).

owner fails to act to evict the possessor, the community may come to rely on the possessor as the person with whom to deal concerning the property.⁶⁵

Professor Carol Rose articulated a related view of the doctrine:

No matter how much the doctrine of adverse possession seems to reward the one who performs useful labor on land at the expense of the lazy owner who does nothing, the crucial element in all these situations is, once again, communication. “Possession” means acts that “apprise the community[,] . . . arrest attention, and put others claiming title upon inquiry.”⁶⁶

Clear title reduces resource-wasting conflict and increases beneficial transactions.⁶⁷

A potential adverse possessor faces risks and rewards. The potential reward for the possessor is not only legal title to the property, but the benefits that come with the use during the time of the adverse possession. The risks involve investment in the property that ultimately may not inure to the benefit of the possessor.⁶⁸ This risk becomes a reality when the owner of the property successfully seeks ejection prior to the expiration of the statute of limitations.

What if we called the users of copyrighted works whose owners cannot be located “adverse possessors”?⁶⁹ Translating the concepts pertinent to adverse possession into the realm of intangible property faces immediate problems. What constitutes actual possession of an intangible asset? One can certainly “use” intangible property, for example by reproducing it or publicly performing it, but “possession” per se is fundamentally difficult. Further, how would an adverse possessor of intangible property demonstrate that their possession was exclusive? By excluding others from real property an adverse possessor appears to the community to be the owner.⁷⁰ The ways to exclude others from using a copyright include threatening litigation, offering licenses, and, ultimately, suing for infringement. Non-owners face serious

65. *Id.* at 79.

66. *Id.* at 80 (citation omitted).

67. *Id.* at 81.

68. It is an open question, based on the particular facts of a specific case, whether a quantum meruit claim would be successful.

69. There have been some copyright cases that have employed the doctrine of adverse possession. *See, e.g.,* *Gee v. CBS, Inc.*, 471 F. Supp. 600, 642–45 (E.D. Pa. 1979). However, recent cases have resoundingly rejected application of the doctrine as state law that is preempted by federal copyright. *Advance Magazine Publishers, Inc. v. Leach*, 466 F. Supp. 2d 628, 634–35 (D. Md. 2006); *see also* *Daus*, *supra* note 59, at 55–56 (1992).

70. *Rose*, *supra* note 64, at 81.

consequences if they engage in any of these activities.⁷¹ Finally, the requirement that the use be “open and notorious” provides the owner of real property with clear notice of the adverse possessor’s action. Yet, open use of a copyrighted work may not be sufficient to provide the copyright owner with notice of the use.⁷²

Setting aside the problematic doctrinal application issues, it is worth considering the application of the underlying policy justifications for adverse possession to the situation facing potential users of a copyrighted work whose owner cannot be located. If adverse possession applied, the neglectful or absentee copyright owner would face a risk of loss of title to the copyright by failing to maintain contact information in accessible directories, such as the Copyright Office.⁷³ Would it be appropriate to reward the user of the copyrighted work with title to the copyrighted work, presumably for the remaining duration of the copyright term?⁷⁴ The productive user of the work would have the potential of benefiting not only from the uses he makes but also from the potential ownership of the copyright after the possessory period expired. Finally, and perhaps most importantly, if such a new owner were identified the world would now have an owner with whom they could transact concerning the copyrighted work. Because so much of the so-called “orphan work problem” stems from a lack of an entity from whom to obtain permission, identifying a new owner solves this roadblock.⁷⁵

The idea of granting rights in a work to someone other than the author may seem antithetical to fundamental aspects of copyright law. While the copyright granted by statute to the author is assignable, it is a very different proposition to grant rights to someone outside of that chain of title. However, as the Supreme Court has recently emphasized, inducing dissemination is part of what Congress can consider when fashioning the

71. Consequences could include liability for fraud and Rule 11 sanctions if an infringement lawsuit is filed. FED. R. CIV. P. 11.

72. Daus, *supra* note 59, at 73–74.

73. I envision that a doctrine of adverse possession would only be applied when a user cannot locate the copyright owner. Certainly checking the records of the Copyright Office would be a logical requirement to impose. This would be similar to the modern statutes that require the adverse possessor to be acting in good faith. *See supra* note 62 and accompanying text.

74. Other durations would be a possibility. For example, a user of a copyrighted work whose owner cannot be located could be rewarded through a statutory grant of exclusive use for a set period of years.

75. France has adopted a model that provides a collective society with the ability to grant licenses for the use of works whose copyright owners cannot be found. *See* Lucie Guibault, *France Solves Its XXe Century Book Problem!*, KLUWER COPYRIGHT BLOG (Apr. 13, 2012), <http://kluwercopyrightblog.com/2012/04/13/france-solves-its-xxe-century-bookproblem/>.

copyright law.⁷⁶ The Constitution requires that exclusive rights be granted to “authors.”⁷⁷ Granting exclusive rights to non-authors, therefore, may be constitutionally impermissible, although Congress could attempt to rest its authority on other Constitutional provisions, such as the Commerce Clause.⁷⁸

More fundamentally, the need to designate someone as the owner in the context of real property has roots in the rivalrous nature of tangible property and a desire to avoid the tragedy of the commons. For copyrights, intangibles that are inherently non-rivalrous, the need to designate an owner stems solely from legal liability rules established by Congress.

The lack of a potential reward of title to the work does not appear to be what is holding back users of copyrighted works whose owners cannot be located. Rather the stumbling block is the potential penalty the user may encounter if the copyright owner surfaces and sues for infringement. Unlike possession of real property, which may result in an eviction action, copyright infringement may result not only in lawsuit seeking an injunction, the arguable equivalent to an eviction, but also a hefty damages award. While an adverse possessor may have to compensate the owner for any proven actual damages to the property, and may even have to disgorge ill-gotten gains from his trespass under a quantum meruit claim, an infringement defendant faces the possibility of an award of statutory damages even if no harm can be shown and no profits were gained through the use.⁷⁹

A final problem in translating the doctrine of adverse possession to copyright is interpreting the Copyright Act’s statute of limitations.⁸⁰ While copyright has a three-year statute of limitation,⁸¹ courts have employed a concept of a “rolling” statute of limitations: each new reproduction, each

76. *Golan v. Holder*, 132 S. Ct. 873, 888 (2012) (citing Thomas B. Nachbar, *Constructing Copyright’s Mythology*, 6 Green Bag 2d 37, 44 (2002)); *see also* *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 555–58 (1985); *Eldred v. Ashcroft*, 537 U.S. 186, 206 (2003); *see generally* Brett M. Frischmann & Mark A. Lemley, *Spillovers*, 107 COLUM. L. REV. 257, 284 (2007) (discussing the incentive for distribution, as well as for creation, of new works).

77. U.S. CONST. art. I, § 8, cl. 8.

78. *See generally* Mark Tushnet, *Constitutional Workarounds*, 87 TEX. L. REV. 1499 (2009) (arguing that some courts have left open the possibility of relying on alternative provisions in the Constitution depending on the nature of the protection afforded); *see, e.g.*, *United States v. Martignon*, 492 F.3d 140, 151 (2d Cir. 2007) (finding criminal anti-bootlegging legislation was not an exercise of the intellectual property power and thus the limitations on that power did not apply).

79. Pamela Samuelson & Tara Wheatland, *Statutory Damages in Copyright Law: A Remedy in Need of Reform*, 51 WM. & MARY L. REV. 439, 480–81 (2009).

80. *Roley v. New World Pictures, Ltd.*, 19 F.3d 479, 481 (9th Cir. 1994).

81. 17 U.S.C. § 507(b) (2010) (requiring a claim be brought within three years from the time that the “claim accrued”).

new distribution of a copy, each new public performance, or each new public display constitutes a new act of infringement for which the copyright owner may pursue an infringement claim.⁸² This view of the statute of limitations doctrine differs from adverse possession, in which initial occupancy or use of land triggers the statute of limitations clock.⁸³ So long as that occupancy or use is open and continuous, the traditional requirements for adverse possession may be met. Unlike with copyright where each new copy distributed starts a new limitations period, each new day of trespassing use does not start a new limitations period running. Rejecting the rolling statute of limitations doctrine in copyright law would be a small but helpful step.⁸⁴

D. PRESCRIPTIVE EASEMENTS

While adverse possession that meets the necessary requirements results in the transfer of ownership of real property in fee simple, prescriptive easements involve non-possessory use of property that results in the creation of an easement, not full title to the property. Called the “first cousin” of adverse possession,⁸⁵ prescriptive easements can arise when an owner of the land fails to make effective objection to another’s use of that land. The elements necessary to establish a prescriptive easement parallel those elements required for adverse possession, and the theoretical justifications are similar. The extent of the use defines the easement that arises.⁸⁶

Similar to easements granted by an owner of land, a prescriptive easement can be either appurtenant or in gross. An easement that is appurtenant has “a dominant estate to which the easement is attached,”⁸⁷ and can be transferred only in connection with the transfer of that dominant

82. “In a case of continuing copyright infringements, an action may be brought for all acts that accrued within the three years preceding the filing of the suit.” *Roley*, 19 F.3d at 481. Section 507(b) “does not provide for a waiver of infringing acts within the limitation period if earlier infringements were discovered and not sued upon.” *Hoey v. Dixel Sys. Corp.*, 716 F. Supp. 222, 223 (E.D. Va. 1989).

83. See H. Tomás Gómez-Arostegui, *Prospective Compensation in Lieu of Final Injunction in Patent and Copyright Cases*, 78 FORDHAM L. REV. 1661, 1712 (2010) (discussing the consequences of labeling a trespass “permanent”).

84. But see Aryeh L. Pomerantz, *Obtaining Copyright Licenses by Prescriptive Easement: A Solution to the Orphan Works Problem*, 50 JURIMETRICS J. 195, 223–27 (2010) (arguing that because the result of adverse possession is the transfer of title, a much longer statute of limitations than the three-year limitations period is necessary).

85. Thomas W. Merrill, *Property Rules, Liability Rules, and Adverse Possession*, 79 NW. U. L. REV. 1122, 1124 (1985).

86. JON W. BRUCE & JAMES W. ELY, JR., *LAW OF EASEMENTS AND LICENSES IN LAND* § 5.4 (2012).

87. 25 Am. Jur. 2d *Easements and Licenses in Real Property* § 10 (2002).

estate. An easement in gross, by contrast, is personal and traditionally cannot be transferred.⁸⁸

If users of copyrighted works whose owners cannot be located were deemed to have obtained a prescriptive easement, their use could continue after the limitations period without fear of infringement liability.⁸⁹ Some of the problems associated with the application of adverse possession elements, as well as its justifications, are similar in the context of prescriptive easements. Although the “use” of the rights of a copyright owner more closely parallels the concepts of use of real property that can lead to a prescriptive easement, the problem of an open use not necessarily informing the owner of the copyright is similar to the problem in the context of adverse possession. The “property” asset of the copyright is not located in one place that the owner can easily monitor.

Additionally prescriptive easements, unlike adverse possession, do not result in a transfer of title to the property being used. Rather, only a license to continue the use is recognized and that license, at least one that is “in gross,” is not transferrable. Thus, the application of the prescriptive easement doctrine does not have the advantage of identifying a “new owner” from whom others interested in making use of the work may seek authorization. Although, as explored above, in the context of intangible property, a new owner is not needed. One interesting aspect of the prescriptive easement may be particularly helpful when considering a user who makes a derivative work incorporating the copyrighted work: perhaps under the right circumstances we should consider the derivative work as the dominant estate to which a prescriptive easement to use the work should attach. In this way a derivative work could continue to be exploited even after the copyright owner resurfaces, so long as the requisite limitations period has passed.

E. DOCTRINE OF WASTE

The doctrine of waste dates back for centuries as one of the ancient writs of the common law. The doctrine of waste “applies when two or more persons have interests in property, but at least one of them is not in possession.”⁹⁰ The doctrine is designed to prevent the person in possession of the property from injuring the non-possessor’s interest in the property. Waste comes in a variety of forms.⁹¹ Permissive waste involves a form of

88. *Id.* § 90.

89. *See* Pomerantz *supra* note 84, at 223.

90. Thomas W. Merrill, *Melms v. Pabst Brewing Co. and the Doctrine of Waste in American Property Law*, 94 MARQ. L. REV. 1055, 1056 (2011).

91. *Id.* at 1057.

nonfeasance, for example a life-tenant who fails to fix a leaking roof that results in damage to the house.⁹² Voluntary waste is a form of misfeasance, for example a life-tenant of an orchard that cuts down the trees.⁹³ Ameliorative waste involves changes to the property that alter the fundamental character of the property but results in a higher value for the property, for example a life-tenant that demolishes a home on the property in order to build a more valuable industrial building.⁹⁴

What if we called the problems associated with copyrighted works whose owners cannot be located a problem of “copyright waste”? One can view the rights of copyright as at least two parties having an interest in the subject property: the copyright owner who has exclusive rights for a period of time and the public who obtains full use of the work upon expiration of the copyright. The copyright owner is the “life tenant,” the one currently “in possession” of the property. The public, as the remainderman, also has an interest in the property. Viewed in this light, a copyright owner who is not available to consider authorization for various exploitation possibilities, possibilities that would result in further dissemination of the work, is engaging in a type of waste. In this context the “waste” produced by the absentee copyright owner is similar to the permissive waste situation exemplified by the life-estate tenant failing to repair the leaky roof. Failing to be available to consider dissemination possibilities results in less distribution and use of the work. Such increased obscurity may ultimately degrade the value of the asset that is the copyrighted work.⁹⁵ Under the doctrine of waste, the public, as the remainderman, has an interest in preventing that waste and in compensation when it has occurred.

If we characterize the situation faced by users of copyrighted works whose owners cannot be located as a “copyright waste problem,” that metaphor might lead us to conclude that different solutions are appropriate. For example, a desire to reduce the waste engaged in by the absentee copyright owners, rather than to protect orphan works or to identify new owners of abandoned works, argues in favor of a solution that lets the public’s interest in greater access prevail.⁹⁶

92. *Id.*

93. *Id.*

94. *Id.*

95. Bernard Lang, *Orphan Works and the Google Book Search Settlement: An International Perspective*, 55 N.Y.L. SCH. L. REV. 111, 116 (2010) (arguing that a “major reason for changing the rules regarding orphan and unregistered works is to revive them as part of our active cultural heritage”).

96. See Herkko Hietanen, *Honey, I Took Out the Trash: Curbside Recycling Motivations and Free Culture Movement*, ISUMMIT08, at 6 (2008), available at <http://commonsresearch.org>.

As Professor John Lovett has highlighted, the doctrine of waste matters most “in settings in which some kind of dramatic, and relatively sudden physical, environmental or economic transformation has taken place.”⁹⁷ That observation rings especially true in copyright today. In two short decades the world experienced the digital revolution. The birth and dramatic expansion of the world wide web itself⁹⁸ created both significant challenges and opportunities for copyright owners and those who are interested in using copyrighted works. During those same two decades, copyright law in the United States has seen the abandonment of formalities⁹⁹ and lengthening of the duration of protection,¹⁰⁰ changes that increased the quantity of copyrighted works whose owners would be difficult, if not impossible, to locate. As Professor Lovett observed, during periods of “radical change,” the “social and economic circumstances affecting the underlying property relationship are changing dramatically.”¹⁰¹ This description fits copyright well. It is during these times of change that courts “confront difficult questions about waste.”¹⁰²

F. HOSTAGE WORKS

The exploration of the different tangible property doctrines demonstrates that alternative metaphors to “orphans” can assist in expanding how we think about the real problems and the possible solutions. All of those tangible property doctrines, however, face challenges translating their policy justifications and their doctrinal elements to the realm of the intangible rights

wdfiles.com/local--files/fc2008-paper3/Honey%20I%20took%20out%20the%20trash.pdf (“Our copyright system, where the majority of works have a low value and transaction costs are high, creates massive amounts of copyright waste”) (citing Edgar J. McCaffery, *Must We Have a Right to Waste*, in *NEW ESSAYS IN THE LEGAL AND POLITICAL THEORY OF PROPERTY* 76 (Stephen R. Munzer ed., 2001)).

97. John A. Lovett, *Doctrines of Waste in a Landscape of Waste*, 72 *MO. L. REV.* 1209, 1212 (2007).

98. Tim Berners-Lee first proposed the hypertext mark-up language that would become the World Wide Web in March 1989. TIM BERNERS-LEE, *WEAVING THE WEB* 21 (2000).

99. In 1989 Congress eliminated the requirement of copyright notice on published copies of a work to maintain federal copyright protection. Berne Convention Implementation Act of 1988, Pub. L. No. 100-568, 102 Stat. 2853 (1988) (codified as amended in scattered sections of 17 U.S.C.). In 1992 Congress eliminated the need to file a renewal registration in order to maintain copyright protection beyond the initial twenty-eight year term for works published prior to January 1, 1978. Copyright Amendments Act of 1992, Pub. L. No. 102-307, 106 Stat. 264 (codified as amended in scattered sections of 17 U.S.C.).

100. 1998 Sonny Bono Copyright Term Extension Act, Tit. I, Pub. L. No. 105-298, 112 Stat. 2827 (1998) (codified as amended in scattered sections of 17 U.S.C.).

101. Lovett, *supra* note 97, at 1012.

102. *Id.*

granted by copyright law. In this Section I provide the alternative metaphor of “hostage works” that I believe better reflects the source of the problem and ultimately facilitates solutions that incorporate lessons learned from the tangible property doctrines.

Copyright law has created a system of automatic rights that lock up the use of expression. The system grants those exclusive rights to authors and subsequent copyright owners with no obligation to identify themselves, or to remain available (let alone an obligation to remain in existence) to consider requests for use of those rights. These works are held hostage by a set of rules that result in an inadvertent lock-up of the expression these works contain.¹⁰³ Whether caused by the lengthy duration of copyright protection, the lack of a registration requirement, the lack of a required maintenance fee to maintain copyright protection, the lack of a notice requirement, the infinite divisibility of copyright ownership, the lack of a requirement for recording transfers of copyright interests, or the exceedingly low threshold for what types of works might be subject to copyright protection, the end result is the same: the copyright owner is difficult or impossible to identify and/or locate. Thus, a market failure results, hindering the public’s access to the expression. Whether caused by the uncertainty of the fair use defense,¹⁰⁴ the hefty statutory damages possible under copyright,¹⁰⁵ or the potential for injunctive relief, the end result is the same: a risk of a magnitude or type of liability that is too great to undertake for many would-be users of hostage works—a market failure results and the public’s access to the expression is hindered. Identifying the problem as a “hostage work problem” therefore seems appropriate.

When viewed as a “hostage work problem” it becomes clear that these works do not need foster parents or protection against inappropriate exploitation—the end result of an orphan metaphor. Nor do these works need new owners—the end result of a metaphor of abandoned or neglected property. What these works need are “special forces” that can free them

103. It is the expression that is locked up because, even during the term of copyright, others are free to use the ideas and concepts embodied in those works. See 17 U.S.C. § 102(b) (2010); see also *supra* note 30.

104. But see Pamela Samuelson, *Unbundling Fair Uses*, 77 *FORDHAM L. REV.* 2537, 2537 (2009) (arguing that the fair use cases are “more coherent and more predictable than many commentators seem to believe”).

105. 17 U.S.C. § 504 (2010) (providing for an award of up to \$30,000 per work infringed, or even higher to \$150,000 per work if the infringement is willful). See also *Rob Reid: The \$8 Billion iPod*, TED (Mar. 2012), http://www.ted.com/talks/lang/en/rob_reid_the_8_billion_ipod.html (video presentation of “Copyright Math (TM),” which vividly demonstrates how quickly statutory damages can add up).

from the constraints placed on them by the combination of the regulatory effects of copyright and the lack of a locatable owner who can grant permission to avoid the consequences of the regulation.

In the previous Section I noted several places where the theoretical underpinnings of property law doctrines do not fit well with copyrights given the inherent characteristics of the rights of copyrights, mainly non-rivalry and non-excludability. The utilitarian justification for copyright protection is premised on a perceived need that to promote progress in knowledge and learning we must guard against not the risk of overuse, but rather the risk of underproduction. However, to achieve the desired progress, the copyright law is designed not only to create an incentive for creation but also an incentive for distribution.¹⁰⁶

In the context of hostage works, the incentive for creation functioned as intended: the work was created. But the incentive for distribution has actually backfired. Instead of a *risk* of underinvestment in distribution, we have a *manifestation* of such underinvestment. Copyright protection is obstructing distribution, not enabling or facilitating it.¹⁰⁷ This is a type of waste: copyright law is “inhibiting access . . . without any countervailing benefit.”¹⁰⁸ In addressing the hostage work problem, we should focus on a solution that reduces the waste by removing the barriers to non-owner distribution.

When one considers the potential users of orphan works, they broadly fall into two categories: access facilitators and derivative work creators.¹⁰⁹

106. See *supra* note 76 and accompanying text. In the sense that the protection afforded by copyright law helps overcome the Arrow’s disclosure paradox, copyright *enables* disclosure and distribution. Arrow’s disclosure paradox describes the problem encountered when one person has developed valuable information but, to commercialize that information, it must be disclosed to others. Once disclosed, there is no reason for the recipient of the information to pay for the information. See Kenneth Arrow, *Economic Welfare and the Allocation of Resources for Invention*, in NATIONAL BUREAU OF ECONOMIC RESEARCH, *THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS* 609, 614–16 (1962).

107. Importantly in the context of hostage works, it is not the copyright owner electing to withdraw the work from public circulation or other exploitation that is interfering with the distribution of the work. See REPORT ON ORPHAN WORKS, *supra* note 2, at 34. Instead it is the inability to locate the owner to seek, and hopefully obtain, the authorization for a desired distribution that creates a hostage work problem.

108. *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 450–51 (1984). The Court specified that this concern was particularly salient in a fair use analysis when there was “no demonstrable effect upon the potential market for, or the value of, the copyrighted work.” *Id.*

109. The Copyright Office Report identified four categories of uses: (1) uses by subsequent creators, (2) large-scale access uses, (3) enthusiast uses, and (4) private uses. REPORT ON ORPHAN WORKS, *supra* note 2, at 36–40. Uses in categories 3 and 4 often can be

Each role has risks and considerations that need to be thought through carefully. The harm that results from the existence of millions of hostage works stems from an inability of the public to have access to those works both in their original form and as incorporated into new derivative works that employ those works.¹¹⁰

The access facilitators are those entities that are interested in distributing copies of the orphan works themselves but fear the infringement liability for doing so. Libraries, archives, museums and other similarly focused entities see providing access to these works as helping advance their core mission of spreading knowledge in their fields. The inability to obtain authorized copies, particularly in desired formats, leads these entities to want to make their own copies. But, without an ability to obtain authorization from the missing copyright owner, engaging in such copying entails a certain amount of risk.¹¹¹ Access facilitators may also include for-profit organizations that would like to harness the commercial value of a storehouse of material and provide additional services that the public may be willing to pay for, such as search engine services or printed copies of works. These types of access facilitators also desire authorization for their copying activity, but again face a risk of infringement liability¹¹² due to an inability to locate the rightful copyright owner. To the extent that the liability risk faced by these actors causes them to forgo providing access to these works, the public loses a benefit it otherwise would have gained had the copyright owner and the access facilitator been able to work out an arrangement.

Derivative work creators include movie producers, educators, authors, musicians, and really anyone who would like to create their own original work of authorship that incorporates the previous work in a way that would

fit into categories 1 or 2 depending on the nature of the use by the private user or enthusiast, e.g., a use that creates a derivative work.

110. Hansen, *supra* note 1, at 8–13.

111. The level of risk is influenced by the problems of narrowly crafted library exemptions, 17 U.S.C. § 108 (2010), the vagaries of the fair use doctrine, 17 U.S.C. § 107 (2010), and the costs of litigating even a strong fair use defense. PRUDENCE ADLER, BRANDON BUTLER, PATRICIA AUFDERHEIDE & PETER JASZI, FAIR USE CHALLENGES IN ACADEMIC AND RESEARCH LIBRARIES 6–7, 13–15 (2010), available at http://www.arl.org/bm~doc/arl_csm_fairusereport.pdf (describing concern among library employees, and discussing practices for minimizing risk of copyright infringement amidst efforts to digitize and preserve collections in non-obsolete formats). The recent willingness of courts to grant attorney's fees to prevailing fair use defendants may begin to have some effect on settlement negotiations. *See, e.g.*, *Brownmark Films LLC v. Comedy Partners*, 101 U.S.P.Q. 2d 1241, 1246–47 (E.D. Wis. 2011).

112. Commercial users face additional risk because their commercial nature may weigh against a claim of fair use. *Sony*, 464 U.S. at 449.

require the copyright owner's authorization.¹¹³ These potential users of orphan works also face a risk of infringement liability if the copyright owner surfaces. Derivative work creators are scared off by the risk of lawsuits from heirs and other claimants coming out of the woodwork if the new use proves economically successful.¹¹⁴ To the extent that the liability risk faced by these actors causes them to forgo the creation and dissemination of the derivative work, the public suffers. The harm here is magnified because not only did the public lose the benefit of access to the hostage work but it also lost out on the benefit of the new work the derivative work creator was going to create but did not.

A solution to the hostage work problem should address both types of users. As I describe below, the access facilitators really are the "special forces" that are freeing the hostages. This role will typically be played by libraries, museums, nonprofit educational institutions, archives, and public broadcasting entities,¹¹⁵ although my proposal is in no way limited to these entities. In addition to the public that will be obtaining access to a work previously held hostage by the rules of copyright, an important beneficiary of the actions of these special forces will be the derivative work creators who should be able to rely on the identification of works as hostage works and, particularly as time passes without a copyright owner surfacing, be willing to use such works.

III. FREEING HOSTAGE WORKS

Freedom for hostage works could come from a reform of the underlying regulations that have led to the hostage crisis in the first place. For example, a return to the duration of copyright protection at the time of the framing of the U.S. Constitution (fourteen years with a possible fourteen-year renewal

113. REPORT ON ORPHAN WORKS, *supra* note 2, at 36 ("The typical scenario might involve an author or publisher that wishes to include a photograph in a new book, or a movie studio that wishes to create a film version of an obscure novel.").

114. The fear is not only of monetary liability but of the injunctive relief that could halt a project in which millions have been invested. *See, e.g.,* Whitmill v. Warner Bros. Entm't, Inc., No. 4:11CV00752 (E.D. Mo. May 24, 2011) (denying requested preliminary injunction by tattoo artist who designed Mike Tyson tattoo that appeared in the movie *Hangover II*), available at <http://docs.justia.com/cases/federal/district-courts/missouri/moedce/4:2011cv00752/113287/41/>.

115. *See, e.g., Final Directive, supra* note 7, at 8 (specifying entities eligible to make certain uses of orphan works are limited to "publicly accessible libraries, educational establishments or museums, as well as by archives, film or audio heritage institutions and public-service broadcasting organisations").

term),¹¹⁶ implementation of a maintenance fee for continued protection,¹¹⁷ or requiring an affirmative act on the part of an author to obtain copyright protection,¹¹⁸ would significantly reduce the quantity of works that become hostages in the first place. Despite the persuasiveness with which reform proposals have been presented,¹¹⁹ I assume that such significant changes will not be forthcoming.¹²⁰ Instead, solving the hostage work problem can come through legislatively or judicially implemented rules that provide protection for responsible “special forces” who act to free hostage works. Also desirable is a robust application of equitable principles for those who use freed works in newly created derivative works.

A. PROTECTING RESPONSIBLE “SPECIAL FORCES”

Freedom for hostage works comes in the form of reliable information concerning the copyright status and the copyright owner of the work. Through the sharing of reliable information,¹²¹ the hostage works will be freed for exposure to interested audiences and potential users of such works. Existing databases can assist with the search for such information,¹²² but someone must invest resources in researching those databases, connecting the discovered information to a particular work, and disseminating the information discovered. Thus, creating incentives to produce and publicize this type of high quality information should be a prime focus of any approach to solving the “hostage work” problem.

In copyright law, as well as in patent law, we are accustomed to granting exclusive rights as an incentive for the creation and distribution of intangible

116. Act of May 31, 1790 § 1, 1 Stat. 124, 124 (repealed 1831). For a thorough discussion of copyright renewal, duration, and term extension, see Tyler T. Ochoa, *Patent and Copyright Term Extension and the Constitution: A Historical Perspective*, 49 J. COPYRIGHT SOC'Y U.S.A. 19, 26–51 (2002).

117. See Lawrence Lessig, *Protecting Mickey Mouse at Art's Expense*, N.Y. TIMES (Jan. 18, 2003), available at <http://www.nytimes.com/2003/01/18/opinion/protecting-mickey-mouse-at-art-s-expense.html> (proposing reintegration of formalities into the Copyright Act).

118. See generally Christopher Sprigman, *Reform(aliz)ing Copyright*, 57 STAN. L. REV. 485 (2004).

119. See, e.g., *id.*; Pamela Samuelson and Members of The CPP, *The Copyright Principles Project: Directions For Reform*, 25 BERKELEY TECH. L.J. 1175, 1222–29 (2010).

120. A variety of obstacles stand in the way of reform, including international treaty obligations. See Sprigman, *supra* note 118, at 545–55.

121. Hansen, *supra* note 3, at 8 (citing *Commission Proposal for a Directive*, *supra* note 7, at 3) (indicating that the proposed EU directive “contemplates that the results of searches would be posted in publicly accessible databases”).

122. See, e.g., *W.A.T.C.H.: Writers Artists & their Copyright Holders*, HARRY RANSOM CENTER, <http://tyler.hrc.utexas.edu/> (last visited June 20, 2012) (providing assistance in determining the status of a particular copyright owner); F.O.B: Firms out of Business, HARRY RANSOM CENTER, <http://tyler.hrc.utexas.edu/fob.cfm> (last visited June 20, 2012).

assets—like information—that we, as a society, believe has value.¹²³ In the context of hostage works, granting an exclusive right to an entity that researches, develops, and publishes information concerning a work could provide an incentive to engage in that activity. The nature of the exclusive right might, for example, permit that entity to control subsequent reproduction of the work. In some ways, this would be similar to the doctrine of adverse possession in which a new owner for the property is designated.¹²⁴ Using a period of exclusivity as a type of “bounty” that would provide an incentive to collect and distribute hostage-work freeing information, however, has the potential to create a second layer of ownership with its own hostage-taking effects. Additionally, granting a new exclusive right for an entity that researches and disseminates this information raises constitutional problems.¹²⁵

More importantly, evidence indicates that there is not a risk of underproduction of this type of information. Recent activities demonstrate that not-for-profit entities are willing to engage in the required information gathering and publishing to free the hostage works so long as they are permitted to distribute copies of the work itself.¹²⁶ For-profit entities, such as Google, likely would engage in this activity as well, so long as there were permissible ways to obtain remuneration.¹²⁷ These not-for-profit and for-

123. For example, in patent law prior to 2003 we encouraged the marshaling of information to challenge questionable pharmaceutical patents by generic manufacturers through the Abbreviated New Drug Application, part of the Hatch-Waxman Amendments. This mechanism granted a 180-day period of semi-exclusivity to “the first generic drug maker who establishes the salability of a product that the name-brand firm’s patent does not control (because the patent is either void or too narrow).” Joseph Scott Miller, *Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents*, 19 BERKELEY TECH. L.J. 667, 723–24 (2004).

124. See discussion *supra* Section II.C.

125. See *supra* note 78 and accompanying text.

126. Many libraries and non-profit entities have demonstrated a desire to be involved in facilitating this access. The libraries that formed the consortium known as the Hathitrust sought to provide access to specifically identified works whose copyright owners the employees of the Hathitrust had researched and could not locate. See *HathiTrust Statement on Authors Guild, Inc. et al. v. HathiTrust et al.*, HATHITRUST DIGITAL LIBRARY (Sept. 15, 2011), http://www.hathitrust.org/authors_guild_lawsuit_statement.

127. While the Google Books project sought to digitize works, Google’s desire to harness the value of the information contained in those works for a variety of purposes could easily translate into a willingness to research the information required to identify hostage works. For example, Google described its efforts in connection with the Google Book Search project and subsequent litigation:

In order to facilitate Rightsholders’ investigations of the copyright status of their books, however, Google put its technical resources to work on the problem. After discussions with Rightsholders and Class Counsel,

profit entities do not appear to need an incentive in the form of an exclusive right. Rather, what these entities need is a reduction in the risk of liability for their actions.¹²⁸

I propose granting immunity from monetary liability for entities that act as responsible “special forces” and free hostage works so long as the entity satisfies two criteria. First, the entity must not be negligent in designating a work as a hostage work or in its approach to correcting status information and removing digital access to a work inaccurately (albeit non-negligently) identified as a hostage. Why negligence is the right standard and what might constitute negligence is explored more fully below. Second, to gain “special forces” immunity from monetary liability, the entity should be required to provide an open access copy of the work with embedded hostage freeing information related to that work.

The requirement of providing open access to the work is a way to ensure the public benefit in return for granting a reduction in liability. In a quid pro quo arrangement that is a familiar way to think about the exclusive rights granted by intellectual property protections,¹²⁹ I suggest that the public benefit is best accomplished by a requirement to provide an open access copy of the work with attached information concerning the copyright and copyright owner status. The details of the open access criteria are explored below.¹³⁰

This proposal for “special forces” immunity could be implemented by judges applying the fair use doctrine: courts should conclude that it is fair use if an entity non-negligently identifies works as hostage works, shares a digital

Google scanned the entirety of the *Catalog of Copyright Entries* from 1923 (before which any books listed in the registry records are in the public domain) to 1978 (after which registrations are listed on the Copyright Office website) and made it freely available for searching through Google Book Search. See *Searching Google's Scans of the Catalog of Copyright Entries*, at <http://books.google.com/googlebooks/copyrightsearch.html>. Thus anyone—both Rightsholders and members of the public—can research the registration status of a pre-1978 book electronically.

Google's Brief in Support of Motion for Final Approval of Amended Settlement Agreement at 60, *Authors Guild v. Google, Inc.*, 770 F. Supp. 2d 666 (S.D.N.Y. 2011).

128. Libraries fear statutory damages, despite statutory provisions that make the imposition of statutory damages against libraries more difficult. See 17 U.S.C. § 504(c)(2) (2010). See also ADLER, ET AL., *supra* note 111, at 3.

129. *Eldred v. Ashcroft*, 537 U.S. 186, 214 (2003) (reviewing what the “bargain” in copyright law entails). See also Shubah Ghosh, *Patents and the Regulatory State: Rethinking the Patent Bargain Metaphor*, 19 BERKELEY TECH. L.J. 1315, 1316–21 (2004) (“Patents are commonly understood as a hypothetical contract between the inventor and the government resulting in a quid pro quo of innovation for exclusivity.”).

130. See *infra* Section III.A.2.

copy of that work under an open access model, and that digital copy contains the hostage freeing information.¹³¹ Similar to the rulings which hold that under certain circumstances reverse engineering of computer software is fair use,¹³² a “special forces” immunity fair use rule, once established, could significantly reduce the perception of risk faced by libraries and other non-profits. One downside of implementing this proposal through the fair use doctrine is that, as an affirmative defense, fair use places the burden on the defendant to prove certain elements in support of the defense.¹³³ In the context of a re-surfacing copyright owner who sues a defendant for use of a work identified as a hostage work, it would seem more appropriate for the plaintiff copyright owner to have to prove the negligence of the defendant, as any claim that relies on negligence as a basis for liability would require.¹³⁴ Under current fair use law, judges could impose this burden on a plaintiff.¹³⁵ Of course, legislative enactment of a type of safe harbor immunity for “special forces” that free hostage works could make clear that the plaintiff would need to prove the defendant’s negligence in order to remove the protection of the safe harbor immunity from monetary liability.¹³⁶

1. *Responsible Identification of Hostage Freeing Information*

To qualify for immunity from monetary liability, an entity that seeks to free hostage works must not be negligent in its identification of a work as a

131. See Jennifer Urban, *Fair Use As a Partial Solution to the Orphan Works Problem*, 27 BERKELEY TECH. L.J. 1379, 1392–1404 (2012).

132. See, e.g., *Sega v. Accoclade*, 977 F.2d 1510, 1527–28 (9th Cir. 1992) (holding that “where disassembly is the only way to gain access to the ideas and functional elements embodied in a copyrighted computer program and where there is a legitimate reason for seeking such access, disassembly is a fair use of the copyrighted work, as a matter of law.”).

133. See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 (1994) (identifying fair use as “an affirmative defense” and noting that defendant would have difficulty prevailing without “favorable evidence about relevant markets”).

134. DAN B. DOBBS, *THE LAW OF TORTS* 269 (2000) (identifying the elements for which the plaintiff bears the burden of proof when demonstrating a prima facie case of negligence).

135. In the *Sony* case the Supreme Court appeared to place a burden on the plaintiff of demonstrating that a particular use was not fair. *Sony Corp. of Am. v. Universal City Studios*, 464 U.S. 417, 451 (1984). Subsequently the Court stated that fair use is “an affirmative defense,” although it did not discuss whether merely the burden of production or the burden of persuasion shifts to the defendant. Therefore it would be entirely appropriate for a court to require a resurfacing copyright owner to carry the burden of persuasion and demonstrate what actions a library, for example, should have taken, under the circumstances, that would have resulted in accurately identifying information that would have led to viable contact with the copyright owner.

136. Congress knows how to create different types of safe-harbors. See, e.g., 17 U.S.C. § 512 (2010) (codifying a set of safe harbor defenses for on-line service providers); 17 U.S.C. § 104A (creating a type of safe harbor for certain actions of “reliance parties” who were using works in the public domain when those copyrights were restored).

hostage work. The entity's negligence would be evaluated when the copyright owner of a work identified as a hostage work sues for copyright infringement.¹³⁷ The doctrine of negligence rests on a duty to exercise due care in one's conduct toward others from which injury may result.¹³⁸ While copyright infringement is often identified as a strict liability offense,¹³⁹ an evaluation of negligence would be entirely appropriate within the fair use defense.¹⁴⁰

Identifying a work as a hostage work has a risk that injury to the copyright owner may result. The injury could result from actions that infringe on the exclusive rights granted to the copyright owner.¹⁴¹ Such infringing conduct would include providing an open access digital copy of the work, a requirement of my proposal discussed in more detail below.¹⁴² The identification of a work as a hostage work should lead to greater freedom for follow-on uses that third parties might engage in after relying on that

137. For any of these issues to be litigated, a copyright owner of an identified hostage work would need to surface and sue. Recently, authors groups asserted claims for infringement against the HathiTrust for its "orphan works project." See discussion of HathiTrust litigation *supra* note 9 and accompanying text. To the extent the claims in the HathiTrust litigation were being asserted by author organizations, rather than specific copyright owners, the court concluded that the domestic author organizations lacked standing under the Copyright Act and thus could not seek to assert the rights of the associations' members. *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939, at *3 (S.D.N.Y. Oct. 10, 2012). The Copyright Act is clear that only those having a legal or beneficial ownership interest in the allegedly infringed copyright may institute an action for infringement. 17 U.S.C. § 501(b). See also *Warren v. Fox Family Worldwide, Inc.*, 328 F.3d 1136, 1145 (9th Cir. 2003) (holding that a composer of a work-for-hire television series soundtrack could not claim beneficial ownership of copyright based on his entitlement to ongoing royalties); *Silvers v. Sony Pictures Entm't, Inc.*, 402 F.3d 881, 890 (9th Cir.) (en banc) (rejecting assertion of beneficial ownership based on assignment of accrued infringement claim without assignment of underlying rights in the copyrighted work), *cert. denied*, 546 U.S. 827 (2005); *Righthaven, LLC v. Hoehn*, 792 F. Supp. 2d 1138 (D. Nev. 2011). The probability of a lawsuit against a user of a hostage work would increase significantly if author trade organizations could instigate litigation.

138. See 57A Am. Jur. 2d *Negligence* § 5 (2012); see also BLACK'S LAW DICTIONARY 712 (9th ed. 2009).

139. *Religious Tech. Ctr. v. Netcom On-Line Comm'n Servs.*, 907 F. Supp. 1361, 1370 (N.D. Cal. 1995).

140. Steven Hetcher, *The Kids Are Alright: Applying A Fault Liability Standard To Amateur Digital Remix*, 62 Fla. L. Rev. 1275, 1279 (2010); Steven Hetcher, *Copytort's Fault Liability Standard*, in COMMON LAW & INTELLECTUAL PROPERTY (Shyamkrishna Balganeshe ed., forthcoming 2012). See also Wendy J. Gordon, *Harmless Use: Gleaning from Fields of Copyrighted Works*, 77 FORDHAM L. REV. 2411 (2009).

141. Note, however, that the "injury" to that intangible right may not, in fact, result in any actual damages that could be proven by the copyright owner.

142. See *infra* Section III.A.2.

identification.¹⁴³ Because those follow-on uses could be infringing, inaccurate identification of a work as a hostage work creates a potential for further injury to the copyright owner by others.¹⁴⁴ Importantly, these injuries only occur if, in fact, there is an existing copyright owner that is interested in continuing to exercise control over the exclusive rights granted by the Copyright Act. However, because there is a risk of injury that can result if a work is inaccurately identified as a hostage work, imposing a duty of care on entities that engage in the identification process would minimize the potential that harm will result.

The Copyright Office's proposal focused on reducing that risk by restricting the remedies available to that later-surfacing copyright owner, so long as the user engaged in a "reasonably diligent search" for the rightsholder.¹⁴⁵ What constitutes a "reasonably diligent search" created anxiety among libraries and archives.¹⁴⁶ A professional negligence standard, in contrast, has a lengthy development in the common law.¹⁴⁷

Professional standards relate to the standard of care typically undertaken by qualified practitioners in the same profession in a given context¹⁴⁸ and will, helpfully, evolve over time as practices evolve.¹⁴⁹ Through the use of a "best practices" model, libraries and archives could establish a standard of care that should be taken in researching the copyright status and copyright owner information for a particular work.¹⁵⁰ In the hostage-freeing context, using a

143. See *infra* Section III.B.

144. Fear of this type of harm appeared to be at the heart of visual artists' objections to U.S. legislation. See *The "Orphan Works" Problem and Proposed Legislation: Hearing Before the H. Subcomm. on Courts, the Internet, and Intellectual Prop., Comm. on the Judiciary*, 110th Cong. (2008) (statement of Marybeth Peters, Register of Copyrights), available at <http://www.copyright.gov/docs/regstat031308.html>.

145. REPORT ON ORPHAN WORKS, *supra* note 2, at 96.

146. *Id.* at 108–09 (discussing the resistance of stakeholders to a formal rulemaking to define the actions that would constitute a diligent search).

147. See Cornelius J. Peck, *Negligence and Liability without Fault in Tort Law*, 46 WASH. L. REV. 225, 226–27 (1971).

148. Dobbs, *supra* note 134, at 633.

149. The ability of the common law to evolve is considered one of its hallmarks. Karl Llewellyn, *THE COMMON LAW TRADITION* 4–7 (1960).

150. The best practices model has worked well in certain fair use areas. See, e.g., CENTER FOR SOCIAL MEDIA, AMERICAN UNIVERSITY, DOCUMENTARY FILMMAKERS' STATEMENT OF BEST PRACTICES IN FAIR USE (2005), available at http://www.centerforsocialmedia.org/sites/default/files/fair_use_final.pdf; ASSOCIATION OF RESEARCH LIBRARIES ET AL., CODE OF BEST PRACTICES IN FAIR USE FOR ACADEMIC AND RESEARCH LIBRARIES (Jan. 2012), available at <http://www.arl.org/bm~doc/code-of-best-practices-fair-use.pdf>. Importantly, the *Best Practices for Documentary Filmmakers* has had an effect on the practices of insurance companies that operate in that market. PATRICIA AUFDERHEIDE & PETER JASZI, *RECLAIMING FAIR USE: HOW TO PUT BALANCE BACK IN COPYRIGHT* 102–05 (2011).

professional negligence framework could address issues such as which databases concerning ownership or entity information should be consulted. Such a framework could also establish whether hostage freeing entities should first provide information concerning works they have identified¹⁵¹ and then wait a specified amount of time until actually making a digital version available.¹⁵²

The duty of care framework could also include standards for sharing particular information that led to a determination that the work is a hostage work, and a standardized format for that information. Importantly, I propose that the information must include the name of the entity making the identification and the date on which that identification is being made public. Finally, the access provider should be required to provide that attached information in both human readable and machine readable language.¹⁵³

2. *Requiring Open-Access to Works with Accompanying Freeing Information*

Establishing a “special forces” immunity must provide a public benefit. Whether that benefit is conceived of as a quid pro quo for the offered immunity from monetary liability, or as a type of compensation for waste caused by the absentee copyright owner,¹⁵⁴ benefitting the public should be primary in the design of a solution to the hostage work problem. Copyright’s

151. Whether the duty of care would impose a requirement to disclose identified hostage works on a centralized database is also something that could evolve as hostage freeing practices matured. The possibility of reliable databases of orphan works was a component of the Orphan Works Act of 2008, which would have tasked the Copyright Office with certifying reliable databases. *See* The Orphan Works Act of 2008, H.R. 5889, 110th Cong. (2008); Shawn-Bentley Orphan Works Act of 2008, S. 2913, 110th Cong. (2008). If such databases evolved, the provenance information attached to the open access copy of the hostage work could indicate in which database the authoritative provenance information is kept. This would facilitate the correction of erroneous designations.

152. For example, the HathiTrust proposed a project for identifying works whose copyright owners could not be located that involved a ninety-day notice period during which time copyright owners could come forward to correct the identification of the work as not having a locatable copyright owner. While no copyright owner came forward, several authors were identified, and HathiTrust nonetheless found itself embroiled in litigation. *See Authors Guild, supra* note 9 and accompanying text. The difference between copyright owners and authors is critical in the context of who can provide authorization for otherwise infringing activity. *See supra* note 25 and accompanying text.

153. The requirement to include this information in both human readable and machine readable is familiar to those using Creative Commons licensing tools. *See RDFa*, CREATIVE COMMONS (Mar. 9, 2012), <http://wiki.creativecommons.org/RDFa> (explaining that by using machine readable information concerning a work, “CC licensed objects can be discovered by search engines and auto-discovery mechanisms without the need for a human to hand-curate content directories or lists”).

154. *See supra* Section II.E.

primary purpose is to benefit the public, specifically to promote the progress of knowledge and learning.¹⁵⁵ Copyright seeks to promote that progress by granting exclusive rights that will provide an incentive for creation and dissemination of new works of authorship.¹⁵⁶ When copyright is hindering dissemination with no beneficial effect on the creation incentive it provides,¹⁵⁷ the public interest in access to extant works is paramount.¹⁵⁸ Additionally, because the increase of hostage works is a direct consequence of the reduction in works that are in the public domain due to longer terms of copyrights and an elimination of formalities to obtain and maintain copyright protection, the public should be compensated through the solution to the hostage work problem.¹⁵⁹

In the context of hostage works, the public benefit would be best served by requiring entities seeking “special forces” immunity to provide an open access copy of the work itself together with attached hostage freeing information for that work. The requirement that an entity must provide an open access copy of the work in order to obtain immunity prevents users of hostage works from exploiting the work for commercial gain with no guarantee of public benefit.¹⁶⁰

For a work to be an “open access” work, a copy of it must be in digital form and accessible through the Internet.¹⁶¹ Additionally, for a work to be an “open access” work it must be free of legal restrictions that would bar the access of the work over the Internet.¹⁶² Open access works must also be free

155. *Golan v. Holder*, 132 S. Ct. 873, 888 (2012).

156. *See* Landes & Posner, *supra* note 41.

157. Randal C. Picker, *Private Digital Libraries and Orphan Works*, 27 BERKELEY TECH. L.J. 1259, 1280–82 (2012) (explaining how an orphan works licensing regime is likely to have minimal effect on most authors at the time that they are deciding to create a work).

158. When a work’s copyright owner cannot be located, it is safe to assume that author of that work is not receiving any compensation for the exploitation of the work. In that context, it is likely that most author’s preference would be to have the work publicly accessible and their name to live on in connection with that work. Lang, *supra* note 95, at 124, 133.

159. *Id.* at 124, 145. *See also* Pamela Samueslon, *Legislative Alternatives to the Google Book Settlement*, 34 COLUM. J.L. & ARTS 697, 720 (2011) (arguing in favor of open access for orphan works as “more consistent with the utilitarian tradition of American copyright law”).

160. In this way, the proposal seeks to avoid one of the fears that the orphan metaphor implies: exploitation of the orphan for purely personal gain. *See supra* Section II.A (discussing the orphan metaphor).

161. Peter Suber, *Creating an Intellectual Commons through Open Access*, in UNDERSTANDING KNOWLEDGE AS A COMMONS: FROM THEORY TO PRACTICE 171 (Charlotte Hess & Elinor Ostrom eds., 2007) (“Open access (OA) is free online access. . . . The physical prerequisites for OA are that a work be digital and reside on an Internet server.”).

162. *Id.*

of technological protection measures and should not require the user to disclose personal information to obtain access to the work.¹⁶³

Open access facilitates authors' and copyright owners' adoption of less restrictive rights that currently attach automatically under the laws of the United States and other maximalist oriented copyright regimes, including those found in the major multilateral treaties of the Berne Convention¹⁶⁴ and the TRIPS Agreement.¹⁶⁵ Open access licenses such as those facilitated by Creative Commons have been described as attempts to minimize waste utilizing optimal licensing regimes.¹⁶⁶

To date, open access has been based on either author consent or public domain status.¹⁶⁷ My proposal would add a third category: hostage work status. To obtain immunity an entity must embed the information that led it to conclude the work is a hostage work. In open access protocols, information concerning the public domain status, referred to as its provenance, includes who declared the work to be public domain.¹⁶⁸ Similarly, the hostage status information must indicate who is making that declaration and the date of the declaration. This type of hostage provenance information will serve important purposes, especially for follow-on creators.¹⁶⁹

The information that is gathered and disclosed by the special forces in connection with the open access copy must not be subject to claims of exclusivity. For example, that data should be released under express

163. Free access means that the "price" of disclosing personal information should also not be permitted.

164. Berne Convention for the Protection of Literary and Artistic Works, Paris Act, July 24, 1971, 25 U.S.T. 1341, 828 U.N.T.S. 221.

165. Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, *available at* http://www.wto.org/english/docs_e/legal_e/27-trips.pdf.

166. *See* Hietanen, *supra* note 96, at 3 (citing Lawrence Lessig, *Re-Crafting a Public Domain*, 18 YALE J.L. & HUMAN. 56 (2006); Lawrence Lessig, *The Creative Commons*, 65 MONT. L. REV. 1 (2004)).

167. *See* Suber, *supra* note 161.

168. The standards for the communication protocols could be established through the W3C Provenance Working Group. *See, e.g., Connection Task Force Informal Report*, W3C (Sept. 29, 2011), http://www.w3.org/2011/prov/wiki/Connection_Task_Force_Informal_Report#Creative_Commons:_Using_Provenance_in_the_Context_of_Sharing_Creative_Works_.28.2A.2A.2A.2A.29 (noting that when it comes to provenance information, "[t]he core statements needed are who licensed, dedicated to the public domain, or marked as being in the public domain, which work, and when").

169. *See infra* Section III.B.

conditions of no assertion of ownership in that information.¹⁷⁰ This type of metadata must be freely and widely available for re-use.¹⁷¹ One way to satisfy this important responsibility would be to employ the Creative Commons Zero Universal Public Domain Dedication, a legal tool developed to make data available without restrictions on reuse.¹⁷²

Requiring public disclosure of the hostage freeing information permits inspection by others, and may lead to the identification of the copyright owner. If that occurs, that copyright owner should then be able to have the provenance information corrected and, if the owner desires, have that work removed from the open access repository. The ability to correct such information along with an obligation of a designating party or an open access database provider to update records could lead to beneficial effects. First, the longer a work is listed without correction, the more likely the hostage designation is an accurate one. Second, different designating entities are likely to develop differing reliability reputations for accuracy. Third, to the extent a follow-on creator desires to invest in the creation of a new work incorporating a designated hostage work, the length of time that has elapsed from designation and the reputation of the designating entity could help facilitate more calculated risk taking and potentially an insurance market to encourage follow-on creation. The risk for follow-on creators could be decreased even further through robust application of a variety of equitable doctrines discussed below.¹⁷³

While the open access copies of an identified hostage work may proliferate, any subsequent commercial user of that work would be wise to check the records of the entity that identified the work as a hostage work in the first place. As proposed, the identity of that entity is a required element

170. The importance of this information being in the public domain parallels the recent move by the Harvard Library to make twelve million catalog records from Harvard's seventy-three libraries available in accordance with its Open Metadata Policy and under a Creative Commons Zero (CC0) public domain license. See Quentin Hardy, *Harvard Releases Big Data for Books*, NEW YORK TIMES: BITS BLOG (Apr. 24, 2012, 9:00 AM), <http://bits.blogs.nytimes.com/2012/04/24/harvard-releases-big-data-for-books/>; *Millions of Harvard Library Catalog Records Publicly Available: Harvard Library Releases Nearly 100% of Its Records*, THE HARVARD LIBRARY TRANSITION NEWS ARCHIVE (Apr. 24, 2012), <http://isites.harvard.edu/icb/icb.do?keyword=k77982&pageid=icb.page498373>.

171. Europeana has also moved to require that the metadata for digitized objects be public domain. See EUROPEANA, THE EUROPEANA LICENSING FRAMEWORK (2012), available at <http://pro.europeana.eu/documents/858566/7f14c82a-f76c-4f4f-b8a7-600d2168a73d>.

172. Creative Commons Legal Code: CC0 1.0 Universal, CREATIVE COMMONS, <http://creativecommons.org/publicdomain/zero/1.0/legalcode> (last visited June 18, 2012); Diane Peters, *Expanding the Public Domain: Part Zero*, CREATIVE COMMONS (Mar. 11, 2009), <http://creativecommons.org/weblog/entry/13304>.

173. See *infra* Section III.B.

of the information that should be attached to the open access copy of the work. Open access uses could be halted following either a notification from the now-contactable copyright owner, or a discovery of relevant contact information as a result of a requirement in the evolving negligence standards for open access providers that would necessitate periodic checks with the designating entity as part of the “duty of care.”¹⁷⁴

B. HOSTAGE WORKS AND DERIVATIVE WORK CREATORS

The second important type of users of hostage works are those that create derivative works. The risk for derivative work creators is that the copyright owner of a hostage work may resurface and seek to enforce its rights against the derivative work creator. Copyright owners have the right to control the preparation of derivative works¹⁷⁵ and can seek both damages and injunctive relief against those who infringe that right.¹⁷⁶ Derivative work creators often have stronger claims to fair use because they are creating a new work that is transformative.¹⁷⁷ If the use at issue is a fair use, the use is not infringing, and there is no hostage work problem—the owner does not need to be contacted because no authorization is needed. The hostage work problem only arises when the use in the derivative work exceeds the bounds permissible under fair use but the copyright owner cannot be located to obtain the desired authorization.

Initially, the difficulty in finding the copyright owner in the first place should weigh heavily in favor of these types of derivative work uses being ruled fair uses, thus expanding the permissible zone of fair use activities.¹⁷⁸ But beyond the fair use inquiry, the hostage work problem for derivative work makers should also trigger a robust application of equitable doctrines. First, courts should be extremely hesitant in granting injunctive relief against

174. The machine readable nature of these records is a critical component that will permit such record updating and periodic checks to occur efficiently. *See supra* note 153.

175. 17 U.S.C. § 106(2) (2010).

176. *Id.* §§ 502, 504.

177. *See* *Campbell v. Acuff-Rose*, 510 U.S. 569, 579 (1994) (“[T]he more transformative the new [derivative] work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use.”). The importance of transformative uses to prevailing on a fair use defense can be seen in many recent fair use cases; *see, e.g.*, *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605 (2d Cir. 2006) (affirming summary judgment in favor of publishers who reproduced images of concert posters throughout a biographical book because presenting them in a new context was sufficiently transformative); *Perfect 10, Inc. v. Amazon, Inc.*, 508 F.3d 1146 (9th Cir. 2007) (vacating a preliminary injunction because a search engine’s display of thumbnail images of copyrighted photographs was a “significant transformative use,” making a fair use defense likely to succeed).

178. *Urban, supra* note 131, at 1394–1401.

derivative work creators who use a work believing in good faith that the work is a hostage work. Second, courts should reject the rolling statute of limitations doctrine, at least in the context of hostage works. Third, courts should engage in a robust application of the doctrine of laches to provide assurances to derivative work makers that viable claims against them will be limited.

1. *Injunctive Relief*

For derivative work makers, the threat of injunctive relief creates serious impediments to using any copyrighted elements for which clearance cannot be obtained. For example, a movie studio may invest tens of millions of dollars creating a movie. If that movie contains an element that might lead to a strategically-timed injunction by a copyright owner, the movie studio will forgo inclusion of that element rather than risk injunctive relief.¹⁷⁹ Of course the threat of injunctive relief can also lead to acquiescence to disproportionate licensing demands by later-surfacing copyright owners.¹⁸⁰

The Supreme Court has been quite adamant in recent years that injunctions in intellectual property cases must be evaluated under traditional principles of equity.¹⁸¹ Specifically, a plaintiff seeking a permanent injunction must demonstrate:

(1) that it has suffered an irreparable injury; (2) that remedies available at law, such as monetary damages, are inadequate to compensate for that injury; (3) that, considering the balance of hardships between the plaintiff and defendant, a remedy in equity is warranted; and (4) that the public interest would not be disserved by a permanent injunction.¹⁸²

While there is some evidence that the Supreme Court's admonitions have fallen on deaf ears in the lower courts,¹⁸³ there is ample room within these equitable considerations for courts to weigh the good faith belief in the

179. REPORT ON ORPHAN WORKS, *supra* note 2, at 37 (noting that “commercial users tend to be highly sensitive to any injunctive relief available to the surfacing copyright owner, especially in situations where an injunction comes at a critical point in the marketing and distribution of the work”).

180. This problem of “holdouts” is a familiar one and not unique to copyright law. See Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1106–07 (1972).

181. *eBay Inc. v. MercExchange, LLC*, 547 U.S. 388 (2006).

182. *Id.* at 391.

183. Jiarui Liu, *Copyright Injunctions After eBay: An Empirical Study*, 16 LEWIS & CLARK L. REV. 215, 228 (2012) (asserting that in the 506 lower court decisions that substantively weighed in on the availability of injunctive relief post *eBay* from May 15, 2006 through June 1, 2010, only 11.3% of the cases even cited the *eBay* decision).

hostage status of the copyrighted work in determining whether to grant injunctive relief.¹⁸⁴ Instead, courts should only be willing to provide a reasonable license fee award to the plaintiff.¹⁸⁵ If courts routinely deny injunctive relief in the context of hostage works, this would likely lead to more reasonable compensation arrangements if copyright owners do surface.

2. *Statute of Limitations*

Courts should also reconsider the application of the statute of limitations in the context of both open access providers and derivative work creators. The Copyright Act provides that “[n]o civil action shall be maintained under the provisions of this title unless it is commenced within three years after the claim accrued.”¹⁸⁶ As described above, the “rolling” statute of limitations doctrine makes it almost impossible to cut off a copyright owner’s ability to sue for ongoing activities that reproduce, distribute, or display the copyrighted work.¹⁸⁷ Eliminating the application of that doctrine in the context of hostage works would be appropriate. If courts determined the limitations period begins when a work is identified as a hostage work and made available in a publicly accessible and searchable repository, the fact that it remains in that repository should not begin a new limitations period each day. The creation of a derivative work should rightly be classified as a “new” act of infringement that would begin a new three-year limitations period, but each new distribution or public performance of that derivative work should not begin a new three-year limitations period.¹⁸⁸

184. The Supreme Court has noted the disservice that injunctive relief might cause in copyright cases. *See, e.g.,* *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578 n.10 (1994) (citing *Abend v. MCA, Inc.*, 863 F. 2d 1465, 1479 (9th Cir 1988), *aff’d sub nom. Stewart v. Abend*, 495 U.S. 207 (1990) (noting the “special circumstances” that would cause “public injury” if the court issued an injunction)).

185. *See* *New York Times v. Tasini*, 533 U.S. 483, 505 (2001) (noting that even once infringement is found, “it hardly follows . . . that an injunction . . . must issue”); *see also* *Davis v. The Gap, Inc.*, 246 F.3d 152 (2nd Cir. 2001) (rejecting claim for \$4 million and holding instead that a reasonable license fee of \$400 to a sunglass designer whose sunglasses were worn by models in a print advertisement was an appropriate remedy).

186. 17 U.S.C. § 507 (2010).

187. *See supra* note 82 and accompanying text.

188. Additionally, the fact that a work was publicly identified as a hostage/orphan by a reliable access provider more than three years prior to the creation of the derivative work could be used in the construction of a laches defense to infringement. It is even possible that derivative work creators would devote resources to assisting access providers in their work or in even becoming access providers themselves. *See infra* Section III.B.3.

Changing the application of the limitations period would encourage prompt publication of hostage freeing information¹⁸⁹ and create an incentive for copyright owners to remain aware of such repositories. The requirement of machine readable metadata would help facilitate automated searching of those repositories.¹⁹⁰ Such open designations with access to digital versions of the work would enrich the public exchange of expressive content.

3. *Laches*

Finally, the application of a robust laches doctrine would create incentives for identifying hostage works early and often. Generally, a “party asserting laches must show: (1) lack of diligence by the party against whom the defense is asserted, and (2) prejudice to the party asserting it.”¹⁹¹ Laches is an equitable doctrine that requires an examination of the facts. As the leading treatise on federal practice provides:

[L]aches does not result from a mere lapse of time but from the fact that, during the lapse of time, changed circumstances inequitably work to the disadvantage or prejudice of another if the claim is now to be enforced. By his negligent delay, the plaintiff may have misled the defendant or others into acting on the assumption that the plaintiff has abandoned his claim, or that he acquiesces in the situation, or changed circumstances may make it more difficult to defend against the claim.¹⁹²

Laches doctrines could come into play when a work is identified as a hostage work and made available in a publicly accessible and searchable repository and no copyright owner has surfaced. A copyright owner who waits to surface until someone else makes a new derivative work should trigger a strong laches defense. As the Ninth Circuit has recognized:

189. The best practices developed could help mediate the tension between desire to publish early and ensuring that best practices are followed in researching the status of the work.

190. The ARROW (Accessible Registries of Rights Information and Orphan Works towards Europeana) project is a good example of various groups working to automate machine readable data to facilitate identifying rightsholders of orphan works in Europe. *About Arrow*, ARROW, <http://www.arrow-net.eu> (last visited June 18, 2012). *See, e.g., Google Alerts*, GOOGLE.COM, <http://www.google.com/alerts> (last visited June 18, 2012) (exemplifying how metadata alerts could be created to allow passive monitoring of newly published orphan works).

191. *Herman Miller, Inc. v. Palazzetti Imps. & Exps., Inc.*, 270 F.3d 298, 320 (6th Cir. 2001).

192. 11A Charles Alan Wright, Arthur R. Miller & Mary Kay Kane, *FEDERAL PRACTICE AND PROCEDURE* 117 (2d ed. 1995) (quoting William Quimby de Funiak, *HANDBOOK OF MODERN EQUITY* 41 (2d ed. 1956)).

It must be obvious to every one familiar with equitable principles that it is inequitable for the owner of a copyright, with full notice of an intended infringement, to stand inactive while the proposed infringer spends large sums of money in its exploitation, and to intervene only when his speculation has proved a success. Delay under such circumstances allows the owner to speculate without risk with the other's money; he cannot possibly lose, and he may win.¹⁹³

Application of the laches doctrine to bar claims against derivative work creators would, over time, reduce the magnitude of the risk faced by these types of hostage work users.

IV. CONCLUSION

Metaphors shape how we approach challenging legal issues. Reorienting the metaphor from “orphans” to “hostages” illuminates the real problem with copyright owners that cannot be located: lock-up of expressive works. Solving the hostage work problem requires creating protections for those who act as “special forces” and free the hostages. If an entity is not negligent in gathering and disclosing information that identifies a work as a “hostage work” and that entity provides an open access copy of the work together with the hostage freeing information, then that entity should be immune from monetary liability for infringement. Copyright owners should retain the ability to obtain injunctive relief to either correct inaccurate status or owner information, or obtain removal of the digital copy of the work from an open access database. This injunctive power would translate into an enforceable obligation of open access providers to update inaccurate information and remove works inappropriately designated as hostage works. For derivative work creators, courts should freely apply equitable doctrines to prevent inappropriate injunctive relief and limit the ability of later re-surfacing copyright owners to sue derivative work creators.

This combination of changes would appropriately reduce the risk that both types of users face when considering utilizing “hostage works” and engaging in activities that would otherwise constitute infringement. At the same time it would provide an appropriate level of protection for copyright owners whose works may be inaccurately identified. In the end, true hostage works would be freed and the public would benefit.

193. *Danjaq LLC v. Sony Corp.*, 263 F.3d 942, 951 (9th Cir. 2001) (quoting *Haas v. Leo Feist, Inc.*, 234 F. 105, 108 (S.D.N.Y. 1916)); *see also* *Chirco v. Crosswinds Cmtys., Inc.*, 474 F.3d 227 (6th Cir. 2007) (holding that laches was properly interposed).

ATOMISM AND AUTOMATION

Molly Shaffer Van Houweling[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1472
II.	COPYRIGHT ATOMISM IN THE DIGITAL AGE	1473
III.	PHOTOGRAPHS AS ATOMISTIC COPYRIGHTED WORKS	1476
IV.	STRATEGIES FOR MITIGATING THE COSTS OF ATOMISTIC OWNERSHIP	1480
	A. REINVIGORATING REGISTRATION.....	1480
	B. DIGITIZING DECENTRALIZED NOTICE.....	1481
V.	ANXIETY ABOUT AUTOMATION	1487
	A. INTEROPERABILITY.....	1487
	1. <i>Technical Interoperability</i>	1487
	2. <i>Legal Interoperability</i>	1490
	B. PERSISTENCE.....	1492
	C. USER COMPREHENSION.....	1496
	D. ACCURACY.....	1497
VI.	CONCLUSION	1500

© 2012 Molly Shaffer Van Houweling. Originally published in the Berkeley Technology Law Journal. Licensed under the Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/> or send a letter to Creative Commons, 444 Castro Street, Suite 900 Mountain View, CA 94041.

[†] Professor of Law, University of California, Berkeley. For helpful discussion and comments on previous drafts, thanks to Robert Barr, Robert Glushko, Robin Kuntz, Mike Linksvayer, Bennett Marks, Deirdre Mulligan, Marianna Murch, Pamela Samuelson, Jason Schultz, Paul Schwartz, Jennifer Urban, and participants in the Berkeley Digital Copyright Project's Symposium on Orphan Works & Mass Digitization: Obstacles and Opportunities. I serve as a Faculty Director of the Berkeley Center for Law and Technology, which supported this research in conjunction with a gift from Nokia (which is a member of the Metadata Working Group, discussed below). I also serve as a member of the Board of Directors of Creative Commons, which is also discussed below. The views I express here—and of course any errors—are my own.

I. INTRODUCTION

Imagine: A budding amateur photojournalist captures the aftermath of a devastating storm using his mobile phone. He then uses the phone to enter a few terms that describe the photo (“hurricane,” “flood”), checks a box to indicate that he wants the photo to be displayed to the public, and checks another to indicate that he is the copyright owner and grants to the public a license to reuse the photo so long as it is attributed to him. He then hits a button to upload the photo wirelessly to the web server of his favorite photo sharing site, where it is immediately available for anyone to view, download, and redistribute via their own favorite photo sharing sites, social networking applications, and news outlets covering the disaster.

Wherever the photo goes on this digital journey, it carries information about its contents, its origin, and its ownership—including the date, time, and location where it was taken; the descriptive terms the photographer added; and his copyright ownership, license, and contact information. When the photo appears online—whether on the photo-sharing site where he originally posted it, or on another photo-sharing site, a social networking platform, a news service, or a peer-to-peer network—this information is communicated to anyone who might want to use the photo yet again. This “metadata” is digitally embedded in the photograph in a way that also makes it possible to find the photo easily by entering the descriptive terms or the photographer’s name into a search engine. The photographer uses this functionality to track the many places online where his photo has been used—with credit to him and his reputation. He receives email inquiries from news outlets who are interested in seeing more of his work.

Technology-savvy photographers know that much of what we have just imagined is not imaginary. Uploading a time-, date-, and place-marked photo from a smartphone to the web while adding descriptive terms that will be displayed along with the photo is becoming commonplace. It is also increasingly commonplace for online photos to be accompanied by copyright licenses that specify how they might be reused. And all of this information can be embedded in digital photo files in a way that can be recognized by subsequent applications and users. But although individual elements of this vision exist, and some have been combined in ways that offer a glimpse of the entire picture, digital photographs do not yet easily and reliably travel on their digital journeys accompanied by information that accurately describes and identifies them and their copyright status. Instead, despite the remarkable potential of digital capture technology and the Internet, digital images can be as mysterious as long-forgotten photos found boxed up in the attic, missing their context and human connections. They become part of our increasingly and frustratingly “atomistic” copyright environment, in which

copyrights are numerous, widely-distributed among often unidentified owners, and fragmented into small and idiosyncratic parts that complicate—or even foreclose—negotiations over reuse of copyrighted works.

In this Article I use digital photographs as a case study of copyright atomism and how automated systems for tagging and tracing might help to alleviate atomism's costs. Part II introduces the copyright atomism concept and explains how atomism has increased in the digital era. Part III focuses on digital photography as an example of extremely atomistic copyright. Part IV describes how embedded metadata might help to alleviate the costs of atomism in this context, but Part V identifies challenges to the metadata solution. Part VI concludes with thoughts about the role of law at the intersection of atomism and automation.

II. COPYRIGHT ATOMISM IN THE DIGITAL AGE¹

Digital technology enables individuals to create and communicate in ways that were previously possible only for well-funded corporate publishers. This makes copyright law newly-relevant to individual amateurs (as opposed to corporate publishers and professional artists and authors) in two different ways. First, individuals are more likely than in the past to reuse copyrighted materials in ways that implicate copyright law—by posting copyrighted photographs on Facebook or other digital platforms without the owner's authorization, for example. Second, individuals are more likely than in the past to use these same digital platforms to publicly distribute *their own* copyrighted creative works. As many observers have noted, most provisions of copyright law in the United States and elsewhere were not designed with these digital-era developments in mind.² Instead, copyright law is too complicated, arcane, and counterintuitive to be well understood by its new targets and beneficiaries.

To be sure, copyright law has changed in recent decades, and some of those changes—specifically, the removal of any formal prerequisites for

1. This Part draws heavily on Molly Shaffer Van Houweling, *Author Autonomy and Atomism in Copyright Law*, 96 VA. L. REV. 549 (2010).

2. Jessica Litman has been an especially insightful observer of this misfit. See Jessica Litman, *Revising Copyright Law for the Information Age*, 75 OR. L. REV. 19, 22–24 (1996); see also Molly Shaffer Van Houweling, *Distributive Values in Copyright*, 83 TEX. L. REV. 1535, 1539 (2005); Pamela Samuelson et al., *The Copyright Principles Project: Directions for Reform*, 25 BERKELEY TECH. L. J. 1175, 1177 (2010) (“Copyright has, of course, always touched and enriched untold ordinary people, as well as specialists, by fostering the provision of entertainment, education, and other information goods and services, but until recently copyright law was relatively invisible to the general public. Amateurs as well as professional artists and authors are now encountering copyright issues on a regular basis.”).

copyright protection—were motivated in part by a desire to make copyright more accessible to individual authors and artists.³ When copyrights are automatically bestowed on individual creators who retain those rights (as opposed to assigning them to intermediaries), the creators have an opportunity to exercise a degree of authorial autonomy that befits the Internet Age. People who once were primarily consumers of creative works now have the capacity to control them.⁴

But these changes in the legal and technological environment have contributed to a contemporary copyright system that is nearly impossible to navigate. When individual creators claim, retain, and manage their own copyrights, they contribute to a troubling phenomenon I have called “copyright atomism”: the proliferation, distribution, and fragmentation of the exclusive rights bestowed by copyright law. An atomistic copyright system is crowded with protected works and rights, owned by rights-holders who are numerous and far-flung, and whose preferences may be idiosyncratic. This situation raises information and transaction costs for participants in a creative marketplace that is full of works to which copyright presumably applies but for which there are no publicly available indications of who owns the copyright, where the owner is, when copyright will expire, and on what terms the works might be licensed.⁵

The history of atomism in copyright, and the broader history and literature about atomistic property rights more generally, suggest a number of different approaches to the challenges that atomism poses. Some of these approaches aim to limit atomism itself—that is, they structure property rights “holistically” so as to restrict, consolidate, unify, and standardize them. An example from copyright law is the work-for-hire doctrine, which avoids the

3. See generally Van Houweling, *supra* note 2, at 1541 n.26 and accompanying text.

4. See generally Van Houweling, *supra* note 1, at 552–53 (observing that “[i]ndividual creators are . . . increasingly harnessing copyright themselves, insisting on ownership of their rights and controlling the ways in which those rights are licensed to others”).

5. The challenges that atomism poses for the digital copyright environment have also been recognized in other property contexts. For example, Robert Ellickson has explained how transaction costs arise in the tangible property context “from the proliferation of boundaries and ownership entities.” Michael Heller’s work has drawn attention to the resource underuse that can arise when property rights fragment into an “anticommons.” Thomas Merrill and Henry Smith focus on the information cost externalities imposed by the proliferation of customized and idiosyncratic property rights. Clarisa Long, Justin Hughes, Henry Smith, and others have applied this transaction and information cost analysis to intellectual property. In my own work, I have examined how these costs may be imposed in the intangible property context by emerging intellectual property licensing practices. For a fuller discussion and citations to the relevant literature, see Van Houweling, *supra* note 1, at 553–54.

consequences of widely-distributed ownership by individual human authors through the legal fiction of employer authorship. This doctrine thus consolidates ownership of contributions to collective projects undertaken in the employment context (and in some cases of works commissioned from independent contractors).⁶

These anti-atomism approaches are often unpopular with individual authors—increasingly so in an age in which the economics of digital creativity make it possible for individuals to create and publish on their own.⁷ More promising in the digital age, then, may be approaches that aim not to maintain a “holistic” (as opposed to atomistic) copyright structure by consolidating ownership, but rather those that address in a more targeted way the information and transaction costs associated with atomism. Copyright history offers examples of this type of approach as well. For example, early copyright law featured systems of registration and other formalities that produced centralized records of copyright ownership, renewal, and transfer—much like the land records that help to facilitate tracing and transactions in the real property context. Contemporary copyright law has largely abandoned this model, however. Registration requirements and other strict “formalities” are disfavored as prerequisites for copyright protection under national copyright laws and international treaties.⁸

Many observers have lamented the demise of formalities in contemporary copyright law and proposed various ways of “re-formalizing” copyright.⁹ These proposals often do not rely on mandatory compliance or on centralized registries, however.¹⁰ Instead, proposals for (and experiments with) other information-based approaches to the problems posed by atomistic copyright are emerging.

One intriguing possibility is that, in place of (or perhaps in conjunction with) centralized registries, information costs could be reduced with machine-readable tags that attach to and travel with digital works. Copyright law has long recognized this idea of decentralized tagging. One of the previously-mandatory formalities under U.S. law was the provision of copyright notice on published works. Indeed, this requirement was more

6. On the work-for-hire doctrine as an anti-atomism device, *see generally* Van Houweling, *supra* note 1, at 557–58, 594–95, 606.

7. *See generally* Van Houweling, *supra* note 1, at 613–21.

8. On the history and contemporary abandonment of copyright formalities, *see* Christopher Sprigman, *Reform(aliz)ing Copyright Law*, 57 STAN. L. REV. 485 (2004).

9. *See, e.g., id.*; Samuelson et al., *supra* note 2.

10. *See id.*

important than registration.¹¹ Today, provision of notice is not required.¹² But contemporary U.S. copyright law does acknowledge, in a more limited way, the value of decentralized notice provision. It incentivizes notice by making additional remedies available where notice has been provided.¹³ And it also forbids (under some circumstances) the removal of copyright tags¹⁴ (a provision that reflects the WIPO Copyright Treaty's protection of "electronic rights management information").¹⁵

This Article begins to explore the technological and legal environment surrounding decentralized information-based interventions in the atomistic copyright system. As a case study, I will focus on the use of copyright-related, machine-readable metadata embedded in digital photographs.

III. PHOTOGRAPHS AS ATOMISTIC COPYRIGHTED WORKS

Photographs are a key element of many of today's most widely-used platforms for user-generated content and social networking. They are typically the work of individual creators, operating independently and retaining ownership of their copyrights.¹⁶ The proliferation of digital

11. In the contemporary period, registration was required under U.S. law only upon renewal. *See generally* Sprigman, *supra* note 8.

12. *See* 17 U.S.C. § 401 (2010) (allowing but not requiring notice).

13. *See id.* §§ 401(d), 504(c).

14. *See id.* § 1202. *See generally* Susuk Lim, *A Survey of the DMCA's Copyright Management Information Protections: The DCMA's CMI Landscape After All Headline News and McClatchey*, 6 WASH. J.L. TECH. & ARTS 297 (2011); Jane C. Ginsburg, *The Right to Claim Authorship in U.S. Copyright and Trademarks Law*, 41 HOUS. L. REV. 263, 283–84 (2004).

15. World Intellectual Property Organization: Copyright Treaty, Dec. 20, 1996, 36 I.L.M. 65 (1997).

16. The most popular platforms (e.g., Flickr, Facebook, Google) specify in their terms of use that contributors retain copyright in their contributions while granting non-exclusive licenses to the platform operators. For Flickr, operated by Yahoo!, see *Yahoo! Terms of Service*, § 9, YAHOO!, <http://info.yahoo.com/legal/us/yahoo/utos/utos-173.html> (last visited June 27, 2012) ("Yahoo! does not claim ownership of Content you submit or make available for inclusion on the Yahoo! Services. However, with respect to Content you submit or make available for inclusion on publicly accessible areas of the Yahoo! Services, you grant Yahoo! the following worldwide, royalty-free and non-exclusive license(s)"); for Facebook, see *Facebook Statement of Rights and Responsibilities*, § 2, FACEBOOK, <http://www.facebook.com/legal/terms> (last visited June 27, 2012) ("You own all of the content and information you post on Facebook For content that is covered by intellectual property rights . . . you grant us a non-exclusive, transferable, sub-licensable, royalty-free, worldwide license"); for Google, including Picasa Web Albums, see *Google Terms of Service*, GOOGLE, <http://www.google.com/intl/en/policies/terms/> (last visited June 27, 2012) ("Some of our Services allow you to submit content. You retain ownership of any intellectual property rights that you hold in that content. . . . When you upload or otherwise submit content to our Services, you give Google (and those we work with) a worldwide license").

photographs with globally scattered individual owners is thus a prime example of atomism in the digital copyright environment. And this atomism has produced concern that information costs could make it difficult to comply with and transact over copyrights in these works.

This anxiety has been most emphatically expressed by professional photographers, who worry that their works will become “orphans” over which they have no effective control.¹⁷ In contrast, other Internet users who share their photographs online do not appear to be motivated by or concerned about copyright and the control it bestows upon them, although copyright nonetheless applies by default to their works. The heterogeneity of both the motivations and the legal and technical sophistication of contributors to this particular cultural environment make it especially difficult to navigate.

Photographs have been recognized as copyrightable subject matter under the U.S. Copyright Act since 1865.¹⁸ Initially there was some question (as there has been in other countries) about whether photographs that aim to capture reality qualify as original works of authorship.¹⁹ But in light of the minimalistic way in which the originality requirement has been interpreted

17. See, e.g., David Riecks, Rivaldo Does & Betsy Reid, *A Metadata Manifesto*, STOCK ARTISTS ALLIANCE (July 2006), <http://www.stockartistsalliance.org/metadata-manifesto-1>.

Lack of information about an image file can delay projects, necessitating additional research to establish licensing rights, obtain clearances, and confirm caption details. This in turn has contributed to the growing problem of misuse of images, whether through error or by intent. Without proper licensing or permissions, users infringe copyright and expose themselves to liability.

Id.

18. See *Goldstein v. California*, 412 U.S. 546, 562 n.17 (1973) (tracing expansion of copyrightable subject matter).

19. See MELVILLE B. NIMMER & DAVID NIMMER, *NIMMER ON COPYRIGHT* § 2.08[E] (2012).

Since its advent in the nineteenth century, photography has posed a basic dilemma for copyright. A rich literature surveys the terrain, discussing whether a photograph simply captures an uncopyrightable fact or qualifies as a work of authorship. The upshot of that debate is that photographs have historically received inferior protection under the copyright laws of other nations Be that foreign treatment as it may, the U.S. experience . . . has embraced photographs as full-fledged members of the copyright genus.

Id. As Nimmer explains, an important early case was *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53 (1884), in which the Supreme Court held that a photograph of Oscar Wilde was copyrightable due to the creative choices (pose, lighting, etc.) made by the photographer. But the Court left open the question of whether “the ordinary production of a photograph” qualifies as original.

under U.S. law, the prevailing view is that almost any photograph “may claim the necessary originality to support a copyright merely by virtue of the photographers’ personal choice of subject matter, angle of photograph, lighting, and determination of the precise time when the photograph is to be taken.”²⁰

Despite this statutory and judicial recognition of their copyrightability, many photographs—especially those by individual amateur photographers—were not subject to copyright protection under U.S. law during photography’s first century. Before the 1976 Copyright Act took effect, publication with notice was a prerequisite for protection that many individual photographers did not satisfy. And even those photos that were initially copyrighted were released into the public domain when their owners failed (as most did) to register their renewals.²¹

This situation changed with the 1976 Act, which granted federal copyright protection to original works upon mere fixation, rather than upon publication with notice.²² In addition, the 1976 Act and subsequent legislation eventually did away with the notice requirement even for published works. And the renewal registration requirement became irrelevant because the dual-term system of protection (initial term subject to extension upon renewal registration) was abandoned and replaced with a unitary term

20. Nimmer & Nimmer, *supra* note 19, § 2.08[E]; *see also* Justin Hughes, *The Photographer’s Copyright—Photograph as Art, Photograph as Database*, 25 HARV. J.L. & TECH. 327, 356 (2012) (“[C]opyright initially struggled with the technology of photography, first came to see the most composed photographs as original art, and then came to reflect artistic theory in accepting that human aesthetic choices permeate many photographs.”); *id.* at 361 (observing the “widespread belief that all photographs are protected by United States copyright law”). *But see id.* at 362 (making the provocative argument that “a large percentage of the world’s photographs are likely not protected by American copyright law because the images lack even a modicum of creativity” and “[i]ndeed, as digitization makes photography more and more ubiquitous, we have probably already crossed a threshold beyond which most of the world’s photographic images are not truly protected by copyright”).

21. *See* William M. Landes & Richard A. Posner, *Indefinitely Renewable Copyright*, 70 U. CHI. L. REV. 471, 505–06 (2003) (documenting renewal rate of approximately 3% for graphic-arts works, including photographs).

22. The 1976 Act introduced the standard that continues to govern today: “Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.” 17 U.S.C. § 102(a) (2010). “A work is ‘fixed’ in a tangible medium of expression when its embodiment in a copy or phonorecord, by or under the authority of the author, is sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated for a period of more than transitory duration.” *Id.* § 101.

calculated based on the life of the author plus fifty (subsequently extended to seventy) years.²³

As the U.S. Copyright Office noted in its 2006 Report on Orphan Works, these developments made it much more difficult to determine the copyright status and owner of protectable works—especially photographs. As the Report explains in its summary of public comments on the topic:

A search for the owner of copyright in a work almost always begins with information available on the work itself. The comments, however, describe numerous situations involving works that bear no information about the author or the owner of copyright in the work—no name of the author, no copyright notice, no title in short, no indicia of ownership on a particular copy of the work *at all*. . . . The comments show that this obstacle is most pervasive—by far—with photographs. Again and again, the comments point to situations involving inadequate information about the author or owner on individual photographs. Numerous individuals complained about situations where they could not use photographs, or did so with trepidation, because they simply had no way of even knowing who took the picture.²⁴

Examples in the Report illustrate that the problem of orphan photographs is not a new problem of the digital era. It has arisen, for example, when people bring copies of old family photos to professional photo finishers, who refuse to make additional copies without proof of copyright ownership or permission.²⁵ The advent of digital photography and web-based photo sharing introduces new challenges, but new solutions as well—as the next Section describes.

23. 17 U.S.C. § 302(a) (2010) (setting duration for works created on or after January 1, 1978, at “a term consisting of the life of the author and 70 years after the author’s death”). See generally Sprigman, *supra* note 8, at 487–88 (documenting removal of formalities requirements under U.S. law).

24. U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 23–24 (2006).

25. *Id.* at 24 (“The most common recurring situation with photographs typically involves the reprinting of old family photographs for preservation, sentimental, or nostalgic purposes. Usually the commenter owns an old, damaged, or deteriorating photograph with no identifying information about the photographer, or outdated information at best. The commenter presents the photographs to a photo-finisher for reproduction, but the finisher refuses to reproduce the work.”).

IV. STRATEGIES FOR MITIGATING THE COSTS OF ATOMISTIC OWNERSHIP

A. REINVIGORATING REGISTRATION

One proposed solution to the orphan works dilemma and other problems posed by atomistic ownership of photographs and other copyrighted works is to reinvigorate the registration system—perhaps with reliance on private registries instead of or in addition to the records maintained by the Copyright Office. For example, the Copyright Principles Project has recommended that “[c]opyright law should encourage copyright owners to register their works so that better information will be available as to who claims copyright ownership in which works,”²⁶ and also that the Copyright Office “should transition away from being the sole registry for copyrighted works and toward certifying the operation of registries operated by third parties, both public and private.”²⁷

The registry solution—whether publicly or privately maintained²⁸—seems less promising for photographs and other images than for some other types of works, however. A text-based registry may not be very useful for someone trying to establish the copyright status of an apparently orphaned photograph, for example. As the Report on Orphan Works explains:

Very often any description of a typical orphan work photograph would be useless: for example, “farmer with three horses, mountains in background.” A user might never match the photograph to the description. Conversely, requiring an owner to monitor a user database would be pointless for the same reason. In addition, adding thumbnail images of the photograph might not help either (even if it were legal and logistically feasible) because the database of images might quickly swell to millions of images: owners and users could still miss each other in the mix.²⁹

Technological advances may make image registries more useful in the future. There have been relevant innovations even in the few years since the Report on Orphan Works was issued. For example, TinEye—a service that launched in May 2008—describes itself as “a reverse image search engine. You can submit an image to TinEye to find out where it came from, how it is being used, if modified versions of the image exist, or to find higher

26. Samuelson et al., *supra* note 2, at 1198 (“Recommendation #1”).

27. *Id.* at 1203 (“Recommendation #2”).

28. See generally Peter S. Menell, *Knowledge Accessibility and Preservation Policy for the Digital Age*, 44 HOUS. L. REV. 1013 (2007) (discussing the importance of public oversight of accessibility and preservation).

29. REPORT ON ORPHAN WORKS, *supra* note 24, at 75 (citations omitted).

resolution versions.”³⁰ TinEye claims to be “the first image search engine on the web to use image identification technology rather than keywords, metadata or watermarks.”³¹ In 2011 Google launched its own reverse image search functionality.³² These services do not purport to provide any information about copyright, but similar technology could be used to make a copyright registry searchable by someone in possession of an image of unknown provenance.

B. DIGITIZING DECENTRALIZED NOTICE

Other technological developments may facilitate a different approach to the information challenges of atomistic copyright in photographs: decentralized notice provision through tagging of the image files themselves (including tagging that refers to a text-based registry or other source of additional information). The Report on Orphan Works notes the importance of the decentralized notice approach:

For authors and copyright owners, marking copies of their works with identifying information is likely the most significant step they can take to avoid the work falling into the orphan works category. This is particularly true for works of visual art, like photographs and illustrations, that otherwise do not contain text or other information that a user can rely on to help determine the identity of the copyright owner.³³

As noted above, this type of marking has long been a part of the U.S. Copyright system, which used to require copyright notice (e.g. “Copyright © 2011 University of California Press”) as a prerequisite for protection of published works.³⁴ Although notice is now voluntary, technological developments may make notice easier for copyright holders to provide, easier for users to find and understand, and therefore more ubiquitous and useful. Relevant technological innovations include techniques and standards for

30. *About*, TINEYE, <http://www.tineye.com/about> (last visited July 31, 2012).

31. *Frequently Asked Questions*, TINEYE, <http://www.tineye.com/faq#what> (last visited June 27, 2012).

32. *Google Search by Image*, GOOGLE, <http://www.google.com/insidesearch/features/images/searchbyimage.html> (last visited June 27, 2012); see also Alex Wild, *Google's Reverse Image Search*, SCIENTIFIC AMERICAN BLOG NETWORK (Aug. 16, 2011), <http://blogs.scientificamerican.com/compound-eye/2011/08/16/googles-reverse-image-search/>.

33. REPORT ON ORPHAN WORKS, *supra* note 24, at 99.

34. See Sprigman, *supra* note 8, at 493–94 (describing evolution of notice requirement under the U.S. Copyright Act).

embedding copyright-related “metadata” in digital files.³⁵ A consortium of digital media companies called the Metadata Working Group³⁶ (“MWG”) is among the advocates for effective use of metadata stored in digital files. The group’s Guidelines for Handling Image Metadata offer this helpful explanation:

Metadata, often referred to as “data about data,” provides interesting information that supplements the primary content of digital documents. Metadata has become a powerful tool to organize and search through the growing libraries of image, audio and video content that users are producing and consuming. This is especially important in the area of digital photography where, despite the increased quality and quantity of sensor elements, it is not currently practical to organize and query images based only on the millions of image pixels. Instead, it is best to use metadata properties that describe what a photo represents and where, when, and how the image was taken.³⁷

The MWG Guidelines identify several key metadata fields, the most familiar of which include:

- keywords (referred to as “tags” in many photo applications)³⁸
- description (often referred to as “caption”)³⁹
- date/time (including when the photo was taken, when it was digitized, and when it was modified)⁴⁰
- location (both location created and location depicted)⁴¹
- creator (or “author”)⁴²
- copyright (specifically, copyright notice and a URL to more information about the copyright).⁴³

35. See generally NAT’L INFO. STANDARDS ORG. UNDERSTANDING METADATA (2004), available at <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>.

36. See METADATA WORKING GROUP, <http://www.metadataworkinggroup.org/> (last visited June 27, 2012) (describing its goals as “[p]reservation and seamless interoperability of digital image metadata” and “[i]nteroperability and availability to all applications, devices, and services”); *Member Companies*, METADATA WORKING GROUP, <http://www.metadataworkinggroup.org/members> (last visited June 27, 2012) (identifying member companies Adobe, Apple, Canon, Microsoft, Nokia, and Sony Corporation).

37. METADATA WORKING GROUP, GUIDELINES FOR HANDLING IMAGE METADATA 8 (Version 2.0, Nov. 2010) [hereinafter MWG GUIDELINES], available at http://www.metadataworkinggroup.org/pdf/mwg_guidance.pdf.

38. *Id.* at 35.

39. *Id.* at 36.

40. *Id.* at 37.

41. *Id.* at 45.

42. *Id.* at 43.

Some of this metadata can be automatically embedded in digital files when they are initially created.⁴⁴ Many modern digital cameras mark photographs with metadata including date, time, and location, for example. Other types of metadata—including copyright information as well as descriptions of the subject matter and other contextual information—can be added manually by the photographer, either photograph-by-photograph or in bulk. Some capture devices (many smartphones, for example) have interfaces that allow manual input of this type of information pre- and/or post-capture. Metadata can also be added using photo-editing software after an image has been loaded onto a computer. And some web-based photo hosting platforms allow users to add additional metadata to hosted photos (during the uploading process and/or afterwards).⁴⁵

The metadata vision that is expressed in the MWG Guidelines is one of interoperability and persistence across the various devices and applications that handle digital images. The ideal is for metadata associated with an image to be durably embedded in the image file in a standards-compliant way that will make it recognizable and usable by subsequent machines, software programs, and ultimately by humans for whom this information is valuable.⁴⁶

In addition to the MWG, organizations representing independent photographers have also taken a leading role in advocating what they see as the proper handling of metadata. In 2006, the Stock Artists Alliance (“SAA”)⁴⁷ released “A Metadata Manifesto,” which insists that “[m]etadata is

43. *Id.* at 42.

44. See Peter Krough, *Metadata Overview*, DPBESTFLOW.ORG, <http://www.dpbestflow.org/metadata/metadata-overview> (last visited June 27, 2012) (including an explanation of “[h]ow does metadata get there?”).

45. See, e.g., *Add Geotags in Picasa*, GOOGLE, <http://support.google.com/picasa/bin/answer.py?hl=en&answer=161869&topic=1689810&ctx=topic> (last visited June 27, 2012).

46. The Guidelines observe—and attempt to rectify—obstacles to this vision. “Different applications and devices have chosen to handle ambiguously or weakly-defined metadata specifications in different ways.” MWG GUIDELINES, *supra* note 37, at 10. “Different applications and devices have chosen to follow different policies in cases where metadata can be stored in more than one standard location.” *Id.* “An application or device often stores proprietary metadata, such as maker notes, within a metadata container. This practice is fragile because such private data can easily be lost when a different application modifies a file.” *Id.* “Some applications and devices usurp general purpose metadata properties to address specific needs. This can cause compatibility problems for applications that correctly use those properties in accordance with the generally accepted specification.” *Id.* “Some applications avoid the complexities of storing metadata within image files altogether and opt, instead, to store it in a separate file or database. This practice can easily result in the loss of metadata when a file is used across several applications.” *Id.*

47. “SAA is an international association of professional photographers and others who create visual works for stock licensing. The mission of SAA is to support and protect the business interests of professional stock photographers worldwide.” *The SAA Mission*

essential to identify and track digital images.”⁴⁸ The document goes on to explain:

Everyone involved with digital images needs to recognize that embedded metadata is an essential part of every digital image. We feel it is the most efficient means of storing information about an image, and the best way to ensure that users can easily find the image and identify its source after it has been distributed or exchanged.⁴⁹

The SAA Manifesto emphasizes, in particular, the role that embedded metadata could play in solving the challenges that atomistic copyright poses both for creators and users—especially in the digital age:

The volume of digital files challenges publishers who need to manage and access them. Busy designers and art directors download preview images to their desktops, only to find weeks or months later they cannot identify the source. Librarians and curators—charged with making more cultural resources available to the public—are already overburdened managing their legacy analog material. Now, they must cope with rapidly expanding digital assets as well. Lack of information about an image file can delay projects, necessitating additional research to establish licensing rights, obtain clearances, and confirm caption details. This in turn has contributed to the growing problem of misuse of images, whether through error or by intent. Without proper licensing or permissions, users infringe copyright and expose themselves to liability. The pressures on image creators—as copyright holders—to protect their intellectual property has intensified since the digitalization and online distribution of their images. If their images cannot be properly identified, they suffer from lost revenues due to missed licensing opportunities.⁵⁰

Among the principles the SAA Manifesto proposes to address these challenges is that “[o]wnership metadata must never be removed.”⁵¹

Similarly, the International Press Telecommunications Council (“IPTC”), a membership organization that develops standards for exchange of content

Statement, STOCK ARTISTS ALLIANCE, <http://www.stockartistsalliance.org/info/about/mission> (last visited June 27, 2012).

48. Riecks et al., *supra* note 17, at 3.

49. *Id.* at 4.

50. *Id.* at 2.

51. *Id.* at 3. The other principles are: “Metadata is essential to identify and track digital images.” *Id.* “Metadata must be written in formats that are understood by all.” *Id.*

(including photographs) within the news industry,⁵² has issued its own “Embedded Metadata Manifesto,” which stresses the importance of persistent and interoperable embedded metadata. The IPTC Manifesto devotes particular attention to copyright-related metadata.⁵³

Like the SAA, the IPTC is concerned about the fate of orphaned photographs in the digital age. Its Manifesto observes that “[p]hotographers, film makers, videographers, illustrators, publishers, advertisers, designers, art directors, picture editors, librarians and curators all share the same problem: struggling to track rapidly expanding collections of digital media assets such as photos and video/film clips.”⁵⁴ To address the copyright aspects of this problem, the IPTC Manifesto insists that “[c]opyright management information metadata must never be removed from the files” because “[o]wnership metadata is the only way to save digital content from being considered orphaned work.”⁵⁵

These manifestos suggest that embedded metadata could address some of the challenges that I associate with atomistic copyright, by reinvigorating the copyright notice function that was long a part of U.S. copyright law. A photographer could embed all of her photos with her name, copyright notice, and a link to license terms and contact information, all in a format that reliably traveled with her digital images across multiple digital platforms—from her camera to her editing software to a web-based platform.⁵⁶ Unlike

52. *IPTC Membership Questions and Answers*, INTERNATIONAL PRESS TELECOMMUNICATIONS COUNCIL (2009), available at http://www.iptc.org/download/public/IPTC-Membership-QandA_2009.1.pdf (“The International Press Telecommunications Council is a consortium of news agencies, newspapers and news system vendors that develops and publishes technical specifications to promote the easy, accurate and inexpensive sharing of news.”).

53. *Embedded Metadata Manifesto (2011)*, INTERNATIONAL PRESS TELECOMMUNICATIONS COUNCIL, [http://www.iptc.org/site/Photo_Metadata/Embedded_Metadata_Manifesto_\(2011\)](http://www.iptc.org/site/Photo_Metadata/Embedded_Metadata_Manifesto_(2011)) (last visited June 27, 2012).

54. *Id.*

55. *Id.*; see also *FAQ: ASMP's Position on Orphan Works*, AMERICAN SOCIETY FOR MEDIA PHOTOGRAPHERS, <http://asmp.org/articles/frequently-asked-questions-about-asmps-position-orphan-works.html> (“[Y]ou should be embedding metadata and communicating with clients to keep that data from being stripped. We are working with software and hardware makers to secure metadata as best we can.”).

56. See generally Abigail L. Dansiger, *Embedded Metadata: Friend or Foe to Our Digital Collections?*, *Library Student Journal* (Jan. 2011), <http://www.librarystudentjournal.org/index.php/ljsj/article/viewArticle/210/290>.

[E]mbedding metadata in images is receiving attention from various parties, such as digital asset managers, professional photographers, federal agencies, and lawyers. These groups have identified embedding metadata as the best way to ensure copyright protection and avoid future litigation

mere textual notices, machine-readable notices could trigger features (pop-up alerts about licensing terms, for example) that would help to ensure that users in fact paid attention and complied with copyright holders' preferences where appropriate. Embedded metadata could also include links to registries that would provide additional information, thus combining decentralized tags with centralized resources like that proposed by the Picture Licensing Universal System ("PLUS"),⁵⁷ which is now testing a registry that it promises will ultimately be used "to find rights and descriptive information . . . for any image, and to find current contact information for related creators, rights holders and institutions."⁵⁸

This vision of a seamless and ubiquitous system of copyright-related (and other) metadata connected to digital image files has not yet been fully realized, however. Indeed, the MWG Guidelines and these metadata manifestos represent attempts to improve metadata-handling practices that are seen as inadequate. In particular, metadata enthusiasts complain that some devices and applications do not embed metadata in a durable and standards-compliant way;⁵⁹ others strip metadata from incoming files;⁶⁰ and many handle metadata in confusing and inconsistent ways that make it difficult for photographers and photo users to understand.⁶¹ Meanwhile, metadata skeptics point out that the best practices for handling embedded

issues or orphan works, which result when copyright holders cannot be identified.

Id.

57. See Riecks et al., *supra* note 17, at 3.

Imagine a world where metadata is ubiquitous. It's a world where images can be easily located and identified by anyone, anywhere. Creators can transmit their images to distributors and users, who instantly integrate these into their systems. Image users can track their digital assets using fully automated systems. A registry—now in development by the Picture Licensing Universal System (PLUS)—will link every image to current information about its source and owner.

Id.

58. *About*, PLUS REGISTRY BETA 1.0, <https://www.plusregistry.org/cgi-bin/WebObjects/PlusDB.woa/1/wo/w8sKt5k98zlrjyCVcb39sg/0.99.27> (last visited July 20, 2012).

59. See *supra* note 46 (noting problems identified by the MWG); see also *Embedded Metadata Manifesto*, *supra* note 53 ("To support the exchange of metadata with content it is a business requirement that file formats embed metadata with the digital file. Other methods like sidecar files are potentially exposed to metadata loss.").

60. See, e.g., Riecks et al., *supra* note 17, at 3 ("Stock photographers, for example, provide digital files to single or multiple distributors who commonly omit metadata from the image files they post online.").

61. MWG GUIDELINES, *supra* note 37, at 10 (citing "significant frustration for users who want consistent metadata interoperability across digital imaging products and services").

metadata are no guarantee that the metadata will be substantively accurate.⁶² In the next Part, I survey and assess these challenges.

V. ANXIETY ABOUT AUTOMATION

As a solution to the problems posed by atomistic copyright in digital photographs, providing distributed digital notice via embedded metadata faces several challenges. These can be roughly categorized as relating to (1) interoperability, (2) persistence, (3) user comprehension, and (4) accuracy. There have recently been important developments on all of these fronts. Progress toward interoperability includes the development and adoption of metadata standards and guidelines for inter-standard compatibility, and the emergence of standardized copyright licenses that increase the potential for legal as well as technical interoperability. Persistence and interoperability are related, as metadata loss can result from incompatibilities. But persistence also depends on policy choices, which I discuss below. As for user comprehension, this too is linked to both technical and legal interoperability, because consumers will find it difficult to understand inconsistent and incompatible schemes. The adoption of standardized licensing tools by major photo hosting services promises to help on this front, while also illustrating the as-yet unrealized potential of license-related metadata. Ensuring the provenance and accuracy of copyright and license-related metadata might be the most intractable challenge—one that technology did not create and may be insufficient to solve completely.

A. INTEROPERABILITY

1. *Technical Interoperability*

Many promising uses of metadata require that it be recognized across multiple devices and applications—for example, from the camera in which an image is captured, to the software with which it is edited, to the web platform on which it is hosted, etc. This type of interoperability relies upon standards for how metadata will be written and processed. As of 2006, an Adobe technical paper lamented that “[b]ecause of the lack of coordination between manufacturers and organizing bodies, much of the potential of metadata remains unrecognized.”⁶³ Adobe has put forward its own proposed solution to this problem in the form of the open-source Extensible Metadata

62. See, e.g., Cory Doctorow, *Metacrap: Putting the Torch to Seven Straw-men of the Metatopia*, THE WELL, <http://www.well.com/~doctorow/metacrap.htm>.

63. *Technical Paper: About Metadata*, ADOBE, http://www.adobe.com/digitalimag/pdfs/phscs2ip_metadata.pdf.

Platform (“XMP”), which the company describes as “a standard for the definition, creation, and processing of metadata.”⁶⁴

Adobe has also partnered with Apple, Canon, Microsoft, Nokia, and Sony to participate in the Metadata Working Group (“MWG”), mentioned above, with the goals of “[p]reservation and seamless interoperability of digital image metadata”; “[i]nteroperability and availability of metadata to *all* applications, devices, and services”; and “[e]nabling emerging metadata-based workflows.”⁶⁵

It remains to be seen whether these efforts will succeed. The MWG membership includes producers of digital photography capture devices and of photo-editing software. Some of the members also maintain web-based services that can be used to share photos.⁶⁶ But it is nonetheless a relatively narrow industry group, not a broad-based standards-setting body. Notably, its membership does not currently include the leading hosts of digital photographs, such as Google, Yahoo! (owner of the online photo-management site Flickr), and Facebook. Google’s Picasa now preserves XMP compatible caption metadata,⁶⁷ and Flickr also preserves at least some metadata on photos uploaded to and downloaded from its service. But neither platform appears to have fully embraced the practices promoted by Adobe and its MWG partners. For its part, Facebook appears to remove almost all metadata from photos uploaded there.⁶⁸

64. *XMP Specification Part I*, at 7 (2010), ADOBE, available at <http://www.adobe.com/content/dam/Adobe/en/devnet/xmp/pdfs/XMPSpecificationPart1.pdf>; see also Fed. Digitization Initiative Still Image Working Grp., *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files* (Aug. 2010), DIGITIZATION GUIDELINES, http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf (“Adobe’s Extensible Metadata Platform (XMP) is another emerging, standardized format for describing where metadata can be stored and how it can be encoded, thus facilitating exchange of metadata across applications.”).

65. MWG GUIDELINES, *supra* note 37, at 11.

66. See, e.g., *PlayMemories Online*, SONY, <http://playmemoriesonline.com/> (last visited Nov. 30, 2012).

67. As Picasa’s user documentation explains: “When you create captions for your photos, they’re saved to your photos and stay with them, even when you upload them, export them, or send them by email. Your captions will be displayed in Picasa, Picasa Web Albums and other EXIF, IPTC and XMP compatible programs.” *Creating and Editing Captions*, GOOGLE (Sept. 27, 2011), <http://picasa.google.com/support/bin/answer.py?hl=en&answer=15055>; see also *Data Stored in Photos (EXIF)*, GOOGLE (Sept. 27, 2011), <http://picasa.google.com/support/bin/answer.py?hl=en&answer=1208119>.

68. See *The Controlled Vocabulary Survey Regarding the Preservation of Photo Metadata by Social Media Websites*, CONTROLLED VOCABULARY, <http://www.controlledvocabulary.com/social-media/> (last visited June 29, 2012).

Where interoperability efforts succeed, users are able to embed metadata with their capture devices and to upload those photos directly to web hosting services where the metadata is preserved and displayed according to the users' preferences. For example, users of image capture devices are increasingly able (1) to elect (now typically via an opt-in mechanism) whether their image will be tagged with geographic information; and (2) to add captions and other descriptive metadata to their images.⁶⁹ Although far from foolproof or ubiquitous, these metadata at least sometimes stay embedded in the images as they are uploaded onto web platforms.⁷⁰

There is no technical obstacle to using the same functionality to embed and preserve metadata related to the copyright and licensing status of photographs.⁷¹ The MWG Guidelines include recommendations for using XMP to embed both a standard copyright notice and a URL to more information about the copyright.⁷² These recommendations promise to promote the inclusion of copyright-related information in standardized locations within image metadata. But they do not do much to promote the standardization of the content of that metadata beyond referring to the basic elements of copyright notice (copyright, date, and owner). For example, if a user wants to provide a URL to more information, what might that information be? And will it be standardized such that subsequent users will understand the terms under which the photograph may be used? Could all of this be simple enough that it could be implemented via the same types of on-screen interfaces that allow smartphone users to specify their geo-tagging preferences for uploaded photos? While software companies and hardware manufacturers focus on technical interoperability, these questions also raise issues of *legal interoperability*.

69. Nokia camera phones feature tools for tagging photos with descriptions and geotags and then for uploading them directly onto either Flickr or Nokia's Ovi platform. *See, e.g., Nokia N900 User Guide, Share Your Images and Videos*, NOKIA, at 76, 82, available at http://www.nokia.com/us-en/support/product/nokia-n900/userguide/?action=onlineuserguidepagechange&pFile=GUID-03220374-A1AE4393-8F2F-D79A29D18A79_FILE001.html (2010). The iPhone Flickr app invites users to add a title, description, tags, and sets. It also asks whether the user wants to share the location as identified by a GPS-enabled phone. The iPhone interface also asks whether the photo should be public or private. *See Flickr for iPhone*, YAHOO!, <http://mobile.yahoo.com/flickr/iphone> (last visited June 29, 2012). Users of Android phones can upload photos directly to Picasa Web Albums and add a caption. Other phones can access the mobile Picasa site via a web browser and upload to it. *See Uploading on Android*, GOOGLE, <http://support.google.com/mobile/bin/answer.py?hl=en&answer=132550>.

70. *See The Controlled Vocabulary Survey*, *supra* note 68.

71. *See* MWG GUIDELINES, *supra* note 37, at 42 (addressing copyright metadata).

72. *Id.*

2. *Legal Interoperability*

One organization working on issues of legal interoperability is Creative Commons.⁷³ This non-profit “that enables the sharing and use of creativity and knowledge through free legal tools”⁷⁴ has promulgated a menu of public licenses that can be associated with copyrighted works (“CC licenses”), permitting reuse of those works under specified circumstances. CC licenses are “written to conform to international treaties governing copyright.” They are “all intended to be effective anywhere in the world, with the same legal effect.”⁷⁵ There are six basic CC licenses applicable to copyrighted photographs. Licensing information that corresponds to these six standardized choices is increasingly becoming a familiar component of web-based photo hosting platforms.

The photo-sharing platform Flickr has offered its users the option of applying CC licenses to their photos since at least 2005. When a Flickr user uploads a photo from her computer, she is invited to add a title, description, and tags. She can create sets and collections in which to organize this and other photos. She can also edit the “owner settings,” including copyright information, specifying either “all rights reserved,” or “some rights reserved.” The “some rights reserved” accompanies images for which the uploading user has elected to apply a CC license to the image, thereby inviting copying and other reuses under specified circumstances. Instead of making this selection for each photo, Flickr users may also select a copyright option as the default for all of their photos. To date, close to two-hundred million Flickr photos have been licensed on CC terms.⁷⁶

More recently, Google’s Picasa Web Albums began to offer the CC licensing option as well. By default, photos uploaded to Picasa Web Albums are displayed with a copyright symbol and the words “all rights reserved.” But this copyright notice can be modified at the option of the user. When viewing her own album using the web interface, a user sees an “edit” link beside the words “All rights reserved.” Clicking “edit” presents the user with the options “Do not allow reuse” or “Allow reuse.” If “Allow reuse” is selected, then the user has the further option of allowing remixing, allowing

73. As acknowledged above, I serve on the Board of Directors of Creative Commons. I also served on its staff from 2001 to 2002.

74. *About*, CREATIVE COMMONS (Aug. 3, 2012, 3:45 PM), <http://creativecommons.org/about>.

75. *CC Affiliate Network: The Licensing Suite*, CREATIVE COMMONS (Mar. 15, 2012, 1:15 PM), http://wiki.creativecommons.org/CC_Affiliate_Network.

76. *Explore/Creative Commons*, FLICKR, <http://www.flickr.com/creativecommons/> (last visited June 29, 2012).

commercial use, and requiring “share-alike” (various combinations of which yield the six basic Creative Commons licenses). Depending on the selection, the photo is subsequently displayed with the words “all rights reserved” or “some rights reserved” (and corresponding icons that represent the CC license choice). A Picasa Web Album user can also apply a Creative Commons license by default to all of her photos by making a selection within the “privacy and permissions” settings of her album.⁷⁷

There remains incompatibility between some of the various CC license types,⁷⁸ and between CC licensing and other terms under which copyright holders might choose to make their photographs available.⁷⁹ But the limited CC menu and its increasing prominence as an option for users of web-based photo platforms promise to deliver to this environment a new degree of legal interoperability—by which I mean standardization of license terms that facilitates comprehension by licensors and licensees, and that facilitates reuse by increasing the probability that multiple resources will be available under the same terms and thus amenable to combined reuse.

Creative Commons aims to link this legal interoperability with metadata interoperability. The licenses are designed to correspond to metadata that can be embedded in digital files, and Creative Commons has developed a metadata schema that is compatible with XMP.⁸⁰ The Picture Licensing Universal System (“PLUS”) has also developed an XMP-compatible “License Data Format,” described as “a metadata ‘schema’ for an embedded image license.”⁸¹

Combining these existing elements would allow photographers to use a simple interface on their capture devices to select copyright licensing options,

77. *See Changing Licensing Options*, GOOGLE (Dec. 11, 2011), <http://support.google.com/picasa/bin/answer.py?hl=en&answer=94293&topic=1689812&ctx=topic>.

78. *Frequently Asked Questions*, CREATIVE COMMONS (Apr. 23, 2012, 6:54 PM), http://wiki.creativecommons.org/Frequently_Asked_Questions#Can_I_combine_two_different_Creative_Commons_licensed_works.3F.

79. *See generally* Van Houweling, *supra* note 1, at 634–35.

80. *XMP*, CREATIVE COMMONS (June 1, 2011, 9:32 PM), <http://wiki.creativecommons.org/XMP>. Based on informal experiments using my own CC-licensed photographs, Flickr and Picasa do not yet appear to embed metadata corresponding to CC license terms—or at least not in ways that are as useful as one would hope. For example, when I downloaded from Flickr a photo that I had posted there with a CC “Attribution” license and then uploaded it to Picasa, its apparent status had reverted to “All Rights Reserved.” Either the license-related metadata was not embedded in the file when I made the selection on Flickr, it was not retained when I uploaded it to Picasa Web Albums, or it was retained but not interpreted and implemented in the way I would expect.

81. *PLUS License Data Format*, PLUS COALITION, INC. (2011), <http://www.useplus.com/useplus/license.asp>.

which could then be embedded into the digital files (along with captions, geographic location, etc.) and retained as those files were loaded onto web-based platforms. Although neither the technical nor legal standards have been universally adopted, and although even adopters have not combined every element of the legal and technical infrastructure I have just described, the progress that has been made illustrates the potential for uniting legally interoperable licensing terms with technically interoperable metadata standards and practices to facilitate the kind of notice provision to which copyright reformers aspire.

B. PERSISTENCE

For metadata to provide useful information about the copyright status, ownership, and license terms of digital images (and other works), it has to stay attached to those images. Photographers who are enthusiastic about the promise of metadata stress this point. For example, the Stock Artists Alliance's 2006 Metadata Manifesto highlights this issue, and its impact on both photographers and image users:

[I]mage creators have no control over what happens to that metadata once their files are circulated. It can be easily removed, as we lack the technological means to create permanent or protected metadata.

Stock photographers, for example, provide digital files to single or multiple distributors who commonly omit metadata from the image files they post online. The problem worsens as files are then forwarded to sub-distributors. Along the way, file names get changed and metadata may be altered or stripped. . . . Typically, there is no metadata in the images, and the stock distributor file names are not helpful for identification. . . .

For image users who are downloading images, archiving, and later repurposing these digital files, the loss of critical information along the way can be detrimental. Without licensing metadata, they cannot determine their rights to use the image. Without contact metadata, they cannot easily inquire about these rights. And without caption metadata, they may not be able to identify who or what is in the image.⁸²

Regarding ownership metadata in particular, the SAA Metadata Manifesto insists that it never be removed without copyright owner authorization:

82. Riecks et al., *supra* note 17, at 3 (emphasis in original).

Ownership metadata must never be removed.

We need to institute standards and best practices in order to protect and preserve critical metadata. Information that identifies the copyright holder must be treated as “read only” or “write once” data, and must never be removed by image distributors and users. The only exception would be changes done with the explicit consent of the copyright owner.

Automated systems for creating and managing digital files need to honor and assist implementation of this principle. Most critically, these systems need to preserve ownership metadata *by default* and discourage removal of other metadata by warning users about the legal implications of removal.⁸³

The Manifesto explains how device interfaces can serve this goal, urging technology providers to “[m]ake preserving metadata the default option when copying, saving or exporting any image” and to “[i]nsure that user interfaces—for operations that remove metadata—are designed to allow users to clearly understand what they are about to do.”⁸⁴

The Metadata Working Group’s Guidelines seem to recognize without fully tackling difficult questions regarding the removal and/or alternation of metadata. The Guidelines articulate rules for “changer applications,” that “read metadata from an image file and then write[] new or modified metadata back to a file.”⁸⁵ Photo sharing platforms like Facebook and Flickr seem to fall within this definition to the extent they modify metadata embedded in uploaded user photos. Section 3.1.2 of the Guidelines addresses removal of metadata by changer applications: “Deletion of metadata **MUST** only be done with specific intent.”⁸⁶ More specifically: “Descriptive metadata, information added by a user, must only be deleted by explicit user intent. Non-descriptive metadata may however be deleted with explicit intent. It should only be modified or deleted if it is known to be inaccurate or problematic.”⁸⁷

This language is unclear on several points. What counts as “explicit user intent” required to delete descriptive metadata? And as for non-descriptive metadata, *whose* explicit intent is required? And *who* should have knowledge that the data is “inaccurate or problematic”? The Guidelines’ Implementation Notes acknowledge this lack of precision:

83. *Id.* at 4 (emphasis in original).

84. *Id.* at 5.

85. MWG GUIDELINES, *supra* note 37, at 16.

86. *Id.*

87. *Id.*

The phrases “done with specific intent” or “by explicit user intent” are intentionally vague. There can be many kinds of intent. For example, a user might ask that a saved file be as small as possible, or be redacted for publication. It is a legitimate design choice for this to involve deletion of metadata. Ideally additional safeguards should be placed around sensitive metadata such as a copyright, but that is the purview of the application, not the MWG.⁸⁸

Although outside the purview of the MWG, this question of when, if ever, copyright information and other metadata should be removed as an image file moves from one platform to another is clearly an important one to some users—and also to a copyright system faced with orphan works and other problems associated with atomistic copyright.

As a legal matter, this type of metadata is protected against removal and alteration under limited circumstances. Provisions of U.S. copyright law added by the Digital Millennium Copyright Act (“DMCA”) address “copyright management information,” which is defined to include specified categories of “information conveyed in connection with copies or phonorecords of a work or performances or displays of a work, including in digital form.”⁸⁹ The covered types of information include, *inter alia*, the author, copyright owner, and terms and conditions for use of the work.⁹⁰

The statute prohibits the provision of false copyright management information and the unauthorized removal or alteration of copyright management information.⁹¹ But both prohibitions also include an intent element. Provision of false information is prohibited when knowing and “with the intent to induce, enable, facilitate, or conceal infringement.”⁹² Removal or alteration is prohibited by someone “knowing, or . . . having reasonable grounds to know, that it will induce, enable, facilitate, or conceal an infringement.”⁹³ This provision seems to recognize the importance of copyright-related metadata and to preserve it against egregious abuses, but it is unlikely that the removal of copyright-related metadata that often occurs when digital images are transferred from one platform to another will violate the prohibition on removal of copyright management information due to the lack of culpable knowledge or intent.

There is, however, another provision of the DMCA that may eventually be more useful for protecting metadata. Section 512 limits the liability of

88. *Id.* at 71.

89. 17 U.S.C. § 1202(c) (2010).

90. *Id.*

91. *Id.* § 1201(a)–(b).

92. *Id.* § 1202.

93. *Id.*

online service providers (“OSPs”) that store material on behalf of their users (and perform other network functions) under specified circumstances.⁹⁴ Widely referred to as the “OSP safe harbor,” this provision imposes various eligibility requirements, including that a covered OSP “accommodates and does not interfere with standard technical measures.”⁹⁵ “Standard technical measures” are defined as:

technical measures that are used by copyright owners to identify or protect copyrighted works and—

(A) have been developed pursuant to a broad consensus of copyright owners and service providers in an open, fair, voluntary, multi-industry standards process;

(B) are available to any person on reasonable and nondiscriminatory terms; and

(C) do not impose substantial costs on service providers or substantial burdens on their systems or networks.⁹⁶

To date, it appears that device manufacturers and software companies, along with some specialized communities of copyright holders and public licensing advocates like Creative Commons, have been most actively involved in developing standards for embedding metadata in digital images and other types of files. This probably does not yet represent the “broad consensus of copyright owners and services providers” to which the statute refers. That standard may be satisfied as these efforts move forward, however.⁹⁷ If it is, social networking services and other platforms used to distribute photographs will have a strong incentive to “accommodate” and “not interfere” with embedded metadata, which for some of them would

94. *Id.* § 512.

95. *Id.* § 512(i)(1)(B).

96. *Id.* § 512(i)(2).

97. See generally Jonathan J. Darrow & Gerald R. Ferrera, *Social Networking Web Sites and the DMCA: A Safe-Harbor From Copyright Infringement Liability or the Perfect Storm?*, 6 NW. J. TECH. & INTELL. PROP. 1, 17 (2007) (“[A]lthough it is likely that watermarks and copyright management systems will *one day* fall into the category of standard technical measures, they do not do so today.”); Lauren G. Gallo, *The (Im)possibility of “Standard Technical Measures” for UGC Websites*, 34 COLUM. J.L. & ARTS 283, 300 (2011) (reviewing the relevant case law and legislative history, noting that “the phrase ‘standard technical measures’ has yet to attach to any existing technology”). Cf. Jane C. Ginsburg, *Separating the Sony Sheep from the Grokster Goats: Reckoning the Future Business Plans of Copyright-Dependent Technology Entrepreneurs*, 50 ARIZ. L. REV. 577, 590–91 & n.57 (2008) (concluding that “[a]rguably, filtering technology might be such a measure,” but that “the definition . . . suggests that the present state of filtering technologies may not suffice, principally because there is not yet an inter-industry consensus regarding the design and implementation of filtering measures”).

mean changing their current practices of removing metadata when uploading user files (unless they could establish that doing so would “impose substantial costs . . . or substantial burdens on their systems or networks”).⁹⁸

C. USER COMPREHENSION

Increased adoption of standards for both the handling of embedded metadata and its content and interpretation will also help to address another challenge: ensuring that users understand how their photos are being marked and how that information can be communicated to others. Some metadata practices have produced user confusion and surprise about both the unexpected retention of metadata and its unexpected removal.

The metadata manifestos quoted in Section V.B, *supra*, reflect the frustration of relatively sophisticated, metadata-savvy photographers who are disappointed when photo metadata is removed without their permission. The opposite type of surprise and confusion has been most notable with regard to privacy-oriented objections to “geo-tagging” of photos with location information that can be automatically collected and embedded into photos by GPS-enabled devices. Depending on how it is subsequently handled, this location information can then be displayed publicly along with the photographs. Photographers who do not understand the capacity of their equipment and the photo-sharing applications they use have been surprised by these revelations.⁹⁹ Increasingly—and likely in response to consumer surprise and dissatisfaction—this functionality is disabled by default and users have to affirmatively choose to turn on geo-tagging functionality.¹⁰⁰

These developments demonstrate the interrelatedness of metadata standardization, user comprehension, and privacy. As metadata practices become more uniform and widespread, users are likely to understand them better (or, in some cases, to object to and prompt changes in emerging practices).

98. 17 U.S.C. § 512(i)(2)(C).

99. See, e.g., Kate Murphy, *Web Photos That Reveal Secrets, Like Where You Live*, N.Y. TIMES, Aug. 11, 2010, at B6, available at <http://www.nytimes.com/2010/08/12/technology/personaltech/12basics.html>.

100. See, e.g., *iOS 4: Understanding Location Services*, APPLE (June 11, 2012), <http://support.apple.com/kb/HT1975>.

Location warnings are the requests made by applications (such as Camera, Compass, and Maps as well as location-dependent third-party apps) to use Location Services with those applications. An application will present a location warning the first time it needs to access Location Services data. Tapping OK will give that app permission to use Location Services as needed.

D. ACCURACY

Some of the leading visionaries of Internet architecture are among the enthusiastic proponents of metadata. Most prominently, Tim Berners-Lee, credited with inventing the World Wide Web,¹⁰¹ articulated his vision for a metadata-enabled “Semantic Web” in 2001:

Most of the Web’s content today is designed for humans to read, not for computer programs to manipulate meaningfully. . . . The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users.¹⁰²

Berners-Lee and the World Wide Web Consortium (“W3C”), which he leads, continue to advocate for the development and adoption of metadata standards required to realize this vision.¹⁰³ But they have met some skepticism, much of it focused on the question of metadata accuracy.¹⁰⁴ Metadata that is perfectly compliant with standards that ensure its interoperability and durability may nonetheless be worse than useless if it is not accurate. To take a simple example: a descriptive tag that identifies a photo of a cat as a photo of a dog will be annoying—not useful—to subsequent collectors of dog photographs. The issue of accuracy is of course even more serious when metadata relates to legal relationships.

How would a would-be re-user of a digital photograph know whether the metadata specifying copyright notice and license terms were embedded by someone who is or has the authority of the copyright holder? Or if, instead, the metadata misidentified the copyright owner and misrepresented the terms under which the true owner was willing to license the work? The short answer is that, given the current state of copyright law in the United States

101. See, e.g., Tim Berners Lee (1955-), BBC HISTORY, http://www.bbc.co.uk/history/historic_figures/berners_lee_tim.shtml (last visited Feb. 6, 2013).

102. Tim Berners-Lee et al., *The Semantic Web*, SCIENTIFIC AMERICAN (May 17, 2001), available at <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>; *Semantic Web Frequently Asked Questions*, W3C (2009), <http://www.w3.org/2001/sw/SW-FAQ>.

103. *Semantic Web*, W3C, <http://www.w3.org/standards/semanticweb/> (last visited Sept. 9, 2012).

104. For an especially pithy example, see Doctorow, *supra* note 62. For an overview of metadata quality issues and existing literature, see Jung-Ran Park, *Metadata Quality in Digital Repositories: A Survey of the Current State of the Art*, 47 CATALOGING & CLASSIFICATION QUARTERLY 213, 225 (2009) (concluding that “the rapidly growing body of digital repositories calls for further investigation of metadata quality” and that “[f]uture studies lie in identifying factors behind inaccurate, inconsistent, and incomplete metadata creation and application”).

and elsewhere, a user cannot answer these questions on the basis of metadata alone. The question of copyright ownership can be complicated and context-specific, depending on facts about the creative process (was the photo taken by the owner of the camera, or by a friend?); about preexisting relationships (was the photographer being employed to take the photo?); and about subsequent events (did he assign his copyright to someone else?).

These problems should give us pause as we consider the feasibility of metadata-based solutions to the problems caused by atomistic copyright. But the problems are not new. The same types of misrepresentations and uncertainties can produce a faulty printed copyright notice in a paper book, or even in the records of the Copyright Office. Since copyrights arise with the fixation of an original work of authorship—and not when notice is provided or when the Copyright Office says so—genuine copyright interests can quite easily coexist with false notices and registrations.¹⁰⁵

Although the basic structure of the Copyright Act contributes to the possibility of conflicts between true copyright ownership and official-seeming representations about ownership, the Act is not entirely inattentive to the problem. If the person who provided the false copyright notice had fraudulent intent, he has committed a crime under U.S. copyright law.¹⁰⁶ Making a knowingly false representation in an application to the Copyright Office is also a crime.¹⁰⁷ The provisions regarding “Copyright Management Information,” added by the DMCA, further prohibit knowingly providing false copyright management information with intent to “induce, enable, facilitate, or conceal infringement.”¹⁰⁸ It seems unlikely, however, that the threat of legal sanction will fully deter (or punish) false provision of copyright notice—especially in the form of digital metadata often supplied

105. See 17 U.S.C. § 411(b)(1) (2010). Knowingly submitting a materially inaccurate application will render the resulting certificate of registration insufficient for purposes of instigating an infringement action, however. *Id.*

106. 17 U.S.C. § 506(c) (2010) (“Any person who, with fraudulent intent, places on any article a notice of copyright or words of the same purport that such person knows to be false, or who, with fraudulent intent, publicly distributes or imports for public distribution any article bearing such notice or words that such person knows to be false, shall be fined not more than \$2,500.”).

107. 17 U.S.C. § 506(e) (“Any person who knowingly makes a false representation of a material fact in the application for copyright registration . . . , or in any written statement filed in connection with the application, shall be fined not more than \$2,500.”).

108. 17 U.S.C. § 1202(a) (2010) (“No person shall knowingly and with the intent to induce, enable, facilitate, or conceal infringement (1) provide copyright management information that is false, or (2) distribute or import for distribution copyright management information that is false.”).

anonymously and with inaccuracies that may result more from carelessness or legal ignorance rather than from fraudulent intent.

Some metadata advocates have proposed technical solutions that address this problem not by directly guaranteeing accuracy or punishing inaccuracy, but rather by facilitating the provision of contextual information that can help users assess the accuracy of metadata and also increase the accountability of metadata providers. For example, in a paper describing the Creative Commons Rights Expression Language (“ccREL”)—an XMP-compatible standard for machine-readable rights-related metadata—Creative Commons identifies the issue of accountability and its special relevance where metadata is embedded in photos and other types of “free-floating content.”¹⁰⁹ In this context, “the method for expressing metadata should facilitate providing publisher accountability at least as strong as the accountability of a Web page with a well-defined host and owner.”¹¹⁰ How would this work? Here is the CC approach:

We handle accountability for free-floating content by connecting any free-floating document to a Web page, and placing the ccREL information on that Web page. Thus, publishers of free-floating content are just as accountable as publishers of Web-based content: rights are always expressed on a Web page. The connection between the Web page and the binary file it describes is achieved using a cryptographic hash, i.e. a fingerprint, of the file. For example, the PDF file of Lawrence Lessig’s “Code v2” will contain a reference to <http://codev2.cc/download+remix>, which itself will contain a reference to the SHA1 hash of the PDF file. The owner of the URL <http://codev2.cc/download+remix> is thus taking responsibility for the ccREL statements it makes about the file.¹¹¹

In short, CC recommends that rights-related metadata include not merely assertions about the copyright owner and license terms applicable to a photo or other piece of content, but also an association with a webpage that can in turn provide a would-be user of the content with some indications of the identity and reliability of the provider of the metadata. This solution is not perfect (as Internet domain name holders can also be difficult to verify and locate), but it does seem to make the accuracy problems associated with digital rights-related metadata akin to the pre-existing threat of false copyright notice on non-digital works. Printed copyright notices and licenses

109. HAL ABELSON, ET AL., CCREL: THE CREATIVE COMMONS RIGHTS EXPRESSION LANGUAGE 11 (Creative Commons 2008), <http://wiki.creativecommons.org/images/d/d6/Ccrel-1.0.pdf>.

110. *Id.*

111. *Id.* at 17.

are not useless in general—and neither is copyrighted-related metadata—merely because it is conceivable that they could be falsely provided by an impossible-to-locate publisher who cannot be held accountable for the falsehood.

In sum, the problem of inaccurate rights-related metadata (like the problem of inaccurate metadata more generally) is a real concern that has rightly drawn the attention of both the advocates and critics of increased use of metadata to manage digital files. But this concern should be evaluated in light of the alternative: files distributed with no copyright information attached to them whatsoever, in a legal environment in which they are automatically protected by copyright. Even if metadata merely give would-be re-users some clues as to the identity of the copyright owner and her preferences—which users may want to confirm through further investigation facilitated by the recommended practice of providing web-based as well as embedded information—in many cases that will be much more information than users would otherwise have had.

Eventually, the accuracy challenge may best be addressed by combining copyright-related metadata with the types of revamped registries that have been proposed by copyright reformers and explored by entities like PLUS and Creative Commons.¹¹² Metadata could both provide a copyright notice and license terms and refer to a centralized registry—perhaps one maintained by a governmental or private body that makes some attempt to verify the ownership claims (akin to the functions provided by land records offices and private title insurance companies in the real estate context).¹¹³

VI. CONCLUSION

This case study of the current practices and prospects for embedding durable copyright-related metadata in digital photographs suggests both the potential and challenges for this type of technological solution to the problems posed by atomism in today's digital copyright environment. We can now imagine how, if the emerging technical and legal standards were combined in a consistent and ubiquitous way, photos that would otherwise be orphaned could be used and reused in compliance with the wishes of their copyright owners—but without costly searches and negotiations. Difficult challenges remain, however, the most intractable of which may be ensuring

112. See *PLUS Registry*, PLUS COALITION, INC., <https://www.plusregistry.org/> (last visited Sept. 9, 2012); Eric Steuer, *Creative Commons Explores a Digital Copyright Registry System*, CREATIVE COMMONS (May 22, 2008), <http://creativecommons.org/press-releases/entry/8306>.

113. See Riecks et al., *supra* note 17, at 3.

metadata accuracy. But this problem is nothing new—either to the metadata community or to copyright law. Indeed, the best way forward may be to combine improving practices for the use and handling of metadata with proposals for revamping copyright registration—using the best of both decentralized and centralized information to provide notice that is interoperable, persistent, comprehensible, and accurate. Such a system could mitigate the costs of atomistic copyright without unduly compromising authorial autonomy and control. Removing the remaining barriers to realizing this vision may be costly and complicated, however. Existing platforms and the software that powers them do not yet incorporate all of the functionality that I have described. Addressing atomism with automation is an attractive vision; but the difficulty and expense of solving copyright challenges with technology should also figure into debates about more fundamental copyright policy reform.¹¹⁴

114. Many of the articles collected in this issue include reform proposals relevant to the problems addressed here, as do, for example, Samuelson et al., *supra* note 2; and Sprigman, *supra* note 8. Most dramatically, perhaps, Justin Hughes has challenged us to rethink the assumption that all photographs in this digital and atomistic age are automatically copyrightable. He suggests that “[a]s we move toward the ubiquitously recorded world in which we are not only constantly photographed, but we ourselves continuously record photographic images and video streams, the whole idea of originality in photography may come under renewed scrutiny.” Hughes, *supra* note 20, at 413.

ORPHAN WORKS AS GRIST FOR THE DATA MILL

Matthew Sag[†]

ABSTRACT

The phenomenon of library digitization in general, and the digitization of so-called “orphan works” in particular, raises many important copyright law questions. However, as this Article explains, correctly understood, there is no orphan works problem for certain kinds of library digitization.

The distinction between expressive and non-expressive works is already well recognized in copyright law as the gatekeeper to copyright protection—novels are protected by copyright, while telephone books and other uncreative compilations of data are not. The same distinction should generally be made in relation to potential acts of infringement. Preserving the functional force of the idea-expression distinction in the digital context requires that copying for purely non-expressive purposes (also referred to as non-consumptive use), such as the automated extraction of data, should not be regarded as infringing.

The non-expressive use of copyrighted works has tremendous potential social value by making search engines possible, and by providing an important data source for research in computational linguistics, automated translation, and natural language processing. Furthermore, the macro-analysis of text is being increasingly used in fields such as the study of literature itself. So long as digitization is confined to data processing applications that do not result in infringing expressive or consumptive uses of individual works, there is no orphan works problem because the exclusive rights of the copyright owner are limited to the expressive elements of their works and the expressive uses of their works.

© 2012 Matthew Sag.

[†] Associate Professor of Law at Loyola University Chicago. Thanks to Pamela Samuelson for encouraging me to pursue this line of inquiry. Thanks also to Jerome Reichman, Matthew Jockers, and the participants at the 2012 Orphan Works and Mass Digitization conference at U.C. Berkeley School of Law. Please address comments to Matthewsag@gmail.com.

TABLE OF CONTENTS

I.	INTRODUCTION.....	1504
II.	UNRAVELING THE DIGITIZATION DEBATE	1506
	A. LIBRARY DIGITIZATION-PRESERVATION	1507
	B. LIBRARY DIGITIZATION-DISTRIBUTION	1508
	C. LIBRARY DIGITIZATION-SEARCH.....	1511
III.	NON-EXPRESSIVE USE	1512
	A. COPYRIGHT, BALANCE, AND THE DISTINCTION BETWEEN IDEAS AND EXPRESSION	1512
	B. NON-EXPRESSIVE USE	1517
	1. <i>Coming to Grips with the Concept of Non-Expressive Use</i>	1517
	2. <i>Examples of the Non-Expressive Use of Expressive Works</i>	1523
	a) Internet Search Engines	1523
	b) Plagiarism Detection Software	1524
	c) Non-Expressive Use and Library Digitization	1525
	C. THE SCOPE OF COPYRIGHT WITH RESPECT TO THE NON- EXPRESSIVE USE OF EXPRESSIVE WORKS	1528
	1. <i>Substantial Similarity</i>	1528
	2. <i>Intermediate Copying</i>	1530
	3. <i>The Implications of Computer Software and Other Functional Works Protected by Copyright Law</i>	1532
	D. ACTIVATING THE PRINCIPLE OF NON-EXPRESSIVE USE THROUGH FAIR USE	1533
	1. <i>Why Fair Use</i>	1533
	2. <i>Application to Fair Use</i>	1535
	a) The “Purpose and Character” of Non- Expressive Uses	1535
	b) Non-Expressive Use and Commercial Fair Use.....	1537
	c) Non-Expressive Use and “Amount and Substantiality”	1538
	d) The Market Effect of Non-Expressive Uses	1539
IV.	CONCLUSION: UNLEASH THE MACHINES.....	1542
	APPENDIX.....	1550

I. INTRODUCTION

Modern technology makes it possible for libraries to scan their paper collections and render them in digital form, making them more useful and

more available than ever before. Modern copyright law ensures that this scanning and digitization process is ensnared in a host of thorny issues.¹ Library digitization² has been rendered thornier still by Google's bold entry into the field in 2004, the ensuing litigation authors and publishers instigated, and the audacity of the class action settlement negotiated in 2008 (and revised in 2009) attempting to resolve that litigation.³

One of the main issues confronting libraries and others with respect to digitization is whether and how to clear rights with respect to works whose copyright owners are not easily found. The existence of these so-called orphan works is one of the most vexing issues in U.S. copyright law today.⁴ One of the main benefits of the class action settlement proposed in relation to *Authors Guild v. Google* was that it would have constituted an expeditious resolution of the orphan works problem standing in the way of library digitization.⁵ However, the treatment of orphan works proposed in the settlement was also one of the primary reasons that the court ultimately rejected it.⁶

This Article aims to untangle the orphan works thicket as it relates to library digitization and show that, correctly understood, *there is no orphan works problem for certain kinds of library digitization*. So long as digitization is confined to data processing applications that do not result in infringing expressive or

1. There is a large literature on library digitization in general and the Google book search litigation. See, e.g., Emily Anne Proskine, *Google's Technicolor Dreamcoat: A Copyright Analysis of the Google Book Search Library Project*, 21 BERKELEY TECH. L.J. 213 (2006); Jonathan Band, *The Long and Winding Road to the Google Books Settlement*, 27 J. MARSHALL REV. INTELL. PROP. L. 227 (2009); James Grimmelman, *D is For Digitize Symposium: An Introduction*, 55 N.Y.L. SCH. L. REV. 11, 12 (2010) (introducing the symposium issue of the *New York Law School Law Review* on the Google Books lawsuit and settlement); Pamela Samuelson, *The Google Book Settlement as Copyright Reform*, 2011 WIS. L. REV. 479 (2011).

2. Library digitization is the process whereby print-based library collections are converted to digital form using scanning and optical character recognition.

3. *Authors Guild et al. v. Google Inc.*, 770 F. Supp. 2d 666, 669–72 (S.D.N.Y. 2011) (reviewing procedural history and rejecting proposed settlement).

4. The U.S. Copyright Office defines “orphan works” as works that are subject to copyright but whose copyright owners “cannot be identified and located by someone who wishes to make use of the work in a manner that requires permission of the copyright owner.” U.S. COPYRIGHT OFFICE, REPORT ON ORPHAN WORKS 1 (2006) available at <http://www.copyright.gov/orphan/orphan-report-full.pdf> [hereinafter REPORT ON ORPHAN WORKS].

5. *Authors Guild*, 770 F. Supp. 2d at 670 (noting that “Older books—particularly out-of-print books, many of which are falling apart buried in library stacks—will be preserved and given new life” (citing Matthew Sag, *The Google Book Settlement & the Fair Use Counterfactual*, 55 N.Y.L. SCH. L. REV. 19, 73 (2010))).

6. *Id.* at 673–86 (rejecting proposed settlement under Rule 23 of the Federal Rules of Civil Procedure).

consumptive uses of individual works, there is no orphan works problem. This conclusion is supported by the fact that copyright owner's exclusive rights are generally limited to the expressive elements of their works and the expressive uses of their works.⁷

II. UNRAVELING THE DIGITIZATION DEBATE

Google entered the world of library digitization in 2004 when it began scanning and digitizing the collections of a number of prestigious private and public academic libraries to make their contents searchable in the same way it makes Internet websites searchable. In many cases, Google also displayed three-line “snippets” of the contents of those books to the general public—just enough to indicate to the searcher whether the text was really responsive to their search term.⁸ Google has been mired in copyright litigation regarding its library digitization project since 2005 when the Authors Guild, along with a group of publishers, sued Google in a class action on behalf of all authors.⁹ Google does not need permission to digitize works in the public domain and the company has also obtained permission from several publishers to include their works in the Google book search engine under agreed terms.¹⁰ However, the company is also digitizing millions of in-copyright works without prior authorization from the relevant copyright owners, and therein lays the core of the dispute.¹¹

The first step towards unraveling the digitization debate is to distinguish between different types of library digitization projects. Google's aspirations for book searches have shifted in a way that complicates the library digitization debate. Initially, the Google Library Project (“GLP”) focused on data processing and search; however, on October 28, 2008, Google, the Authors Guild, and a group of leading publishers proposed a class action settlement that, among other things, would have transformed the GLP into a

7. I first made this argument in an article addressing the significance of copy-reliant technology more generally. This Article refines and extends my earlier analysis. See Matthew Sag, *Copyright and Copy-Reliant Technology*, 103 NW. U. L. REV. 1607 (2009).

8. *Id.* at 1620–22 (describing the operation of the Google book search engine).

9. *Authors Guild*, 770 F. Supp. 2d at 670.

10. Google's Partner Program enables rights owners to opt into book search and allows them to control how their works are searched and displayed. Google has signed up over 20,000 rights holders to this Partner Program. See *Information for Authors and Publishers*, GOOGLE BOOKS, <http://www.google.com/googlebooks/publishers.html>.

11. As of March 2011, Google had scanned at least twelve million books. See *Authors Guild*, 770 F. Supp. 2d at 670.

general distribution platform for electronic versions of books.¹² For the sake of clarity, this Article will refer to the former as “GLP-search” and the latter as “GLP-distribution.” The distinction is important because, although GLP-search has a strong claim to legality under current U.S. copyright law, GLP-distribution does not.¹³

Looking beyond Google, it is useful to think of all library digitization initiatives in three conceptually distinct genres corresponding to the three objectives of library digitization: (1) preserving existing volumes (“library digitization-preservation”); (2) facilitating data analysis and digital searching (“library digitization-search”); and (3) facilitating access to electronic versions of books (“library digitization-distribution”). The legal issues relating to each of these genres must be considered separately.

A. LIBRARY DIGITIZATION-PRESERVATION

Although libraries have certain privileges under the Copyright Act, nothing in the statute expressly allows wholesale library digitization with the exclusive aim of preserving existing volumes. Section 108 of the Copyright Act allows libraries to reproduce and distribute works “for purposes of preservation and security or for deposit for research use in another library” or to replace copies that are “damaged, deteriorating, lost, or stolen,” or for which the existing format has become obsolete.¹⁴ The scope of § 108 is very narrowly tailored and the provision does not authorize a general digitization program for preservation purposes.¹⁵ For example, § 108(b) allows a library to make three copies of any unpublished work in its collection for preservation and security purposes.¹⁶ Section 108(c) also permits a library to

12. Plaintiff’s Motion for Preliminary Settlement Approval, *Authors Guild v. Google Inc.*, 93 U.S.P.Q.2d 1159 (S.D.N.Y. Oct. 28, 2008) (No. 05 Civ. 8136) [hereinafter Motion for Preliminary Settlement Approval]. Settlement negotiations apparently began in the fall of 2006. In response to significant public criticism, including from the Department of Justice, the parties proposed an Amended Settlement Agreement on November 13, 2009. *Authors Guild*, 770 F. Supp. 2d at 671–72.

13. See *infra* Section II.C.

14. 17 U.S.C. § 108(a)–(c) (2010).

15. See Lois F. Wasoff, *If Mass Digitization Is the Problem, Is Legislation the Solution? Some Practical Considerations Related to Copyright*, 34 COLUM. J.L. & ARTS 731, 738 (2011). Wasoff noted:

Current U.S. copyright law has no provision permitting libraries to make preservation copies of published works. Preservation copies are limited to unpublished works; replacement copies can be made of published works if the work is damaged or lost, but only if an unused copy cannot be located at a fair price.

Id.

16. 17 U.S.C. § 108(b).

make three copies of published works to replace a work in the library's collection that is (or was) damaged, deteriorating, lost, or stolen—but only if the library is unable to obtain a new copy at a fair price.¹⁷

The recommendations of the § 108 Study Group and the Copyright Principles Project to expand and clarify the scope of § 108, with respect to preservation, have much to recommend. However, the legal status of digitization aimed solely at preservation is an issue at the periphery of the debate.¹⁸ Even if library digitization-preservation were clearly protected under the Copyright Act, there would still be considerable pressure to address the issues of search and distribution.

B. LIBRARY DIGITIZATION-DISTRIBUTION

In general, the digitization of library books to enable substantial display and/or distribution of e-books clearly implicates the copyright owner's rights. To scan a book is to reproduce the work in a digital copy,¹⁹ and substantial textual displays and distribution of further copies clearly have the potential to substitute for the copyright owner's authorized copies and would not generally be protected by fair use. It is certainly arguable that fair use would protect the display of works that are out-of-print and whose copyright owner or owners cannot be located with reasonable efforts.²⁰ But putting the orphan works issue to one side for the time being, without additional facts, there is nothing to indicate that merely making a work more available is a transformative use that imbues the original "with a further purpose or different character, altering the first with new expression, meaning, or message."²¹ There may be specific instances where such display or distribution would be justified as fair use, or would be protected by the § 108

17. 17 U.S.C. § 108(c). See also Laura N. Gasaway, *Values Conflict in the Digital Environment: Librarians Versus Copyright Holders*, 24 COLUM. J.L. & ARTS 115, 121–23 (2001); Samuelson, *Google Books*, *supra* note 1.

18. SECTION 108 STUDY GROUP, U.S. COPYRIGHT OFFICE, THE SECTION 108 STUDY GROUP REPORT iii–x (2008) (recommending numerous changes to library and archival privilege), available at <http://www.section108.gov/docs/Sec108StudyGroupReport.pdf>; see also Pamela Samuelson et al., *The Copyright Principles Project: Directions for Reform*, 25 BERKELEY TECH. L.J. 1175 (2010) (recommending some updates to library and archival privilege).

19. See 17 U.S.C. § 106(1) (providing that the copyright owner has the exclusive right to reproduce the work in copies).

20. To expand upon my view of this issue would be distracting in the context of this Article.

21. See *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994).

library privilege or some other exception under the Copyright Act—but these would be exceptions to the usual rule.²²

Consider, for example, two features of the GLP provided for in the Amended Settlement Agreement (“ASA”). Unless the rights holder opted out, the ASA would have allowed Google to sell online access to entire books as consumer purchases or “institutional subscriptions.”²³ The ASA also envisaged a default book display of up to 20% of a book, not just a three-line snippet.²⁴ Such extensive displays may well benefit copyright owners by stimulating interest in the entire work, but they also potentially substitute for the original works.²⁵ The ASA would have allowed Google to sell access to copyrighted works in a format and to an extent that substituted for purchase of copyright owner authorized copies. Such an action is well beyond the conceivable parameters of the idea-expression distinction or fair use.²⁶

To many, the legal obstacles confronting a full-fledged e-distribution model of library digitization highlight the failure of the law to adapt to new technology. GLP-distribution, as proposed under the ASA, has been described as “one of the most important applications of digital information technology in the information age.”²⁷ Many out-of-print books are currently available only to those with access to large research libraries. Furthermore, library digitization has the potential to democratize access to these works and create an important sphere of equality of opportunity. If digitization were linked to some kind of payment mechanism it would help authors “breathe

22. A recent empirical study of fair use concludes that transformative use by the defendant is a robust predictor of a finding of fair use; the amount and substantiality of the defendant’s unauthorized use of the plaintiff’s work is a significant factor in litigated fair use cases; but also notes that there is “no evidence that commercial use (in contrast to direct commercial use) reduces the defendant’s chance of maintaining a fair use defense.” Matthew Sag, *Predicting Fair Use*, 73 OHIO ST. L.J. 47, 85 (2012).

23. Amended Settlement Agreement § 4.1, *Authors Guild et al. v. Google Inc.*, 770 F. Supp. 2d 666 (S.D.N.Y. 2011) (No. 05 Civ. 8136) [hereinafter ASA].

24. *Id.* § 4.3(b)(1).

25. *A&M Records v. Napster, Inc.*, 239 F.3d 1004, 1018 (9th Cir. 2001) (holding that increased sales of copyrighted material attributable to unauthorized use should not deprive the copyright holder of the right to license the material).

26. *See Authors Guild*, 770 F. Supp. 2d at 678 (“Google did not scan the books to make them available for purchase, and, indeed, Google would have no colorable defense to a claim of infringement based on the unauthorized copying and selling or other exploitation of entire copyrighted books.”).

27. *See Lateef Mtima & Steven D. Jamar, Fulfilling the Copyright Social Justice Promise: Digitizing Textual Information*, 55 N.Y.L. SCH. L. REV. 77, 104 (2010).

new life into older, out-of-print books that are generally inaccessible to the public and have stopped generating revenue.”²⁸

Copyright law poses an obstacle to the electronic distribution of out-of-print books because of the high costs of proactively clearing rights with copyright owners. As time progresses, things like assignments, deaths, bankruptcies, mergers, spin-offs, asset sales, reversion clauses in publishing contracts, poor private record keeping, and poor public record keeping by the Copyright Office can complicate the question of who owns the work.²⁹ The more time elapses, the higher the likelihood that the public record no longer provides enough information to know whom to ask for permission to use the copyrighted material.³⁰ Changes in copyright law over the years have exacerbated this problem by making the vesting (and continuation) of copyright automatic and by increasing the term of protection to the author’s life plus seventy years, or ninety-five years from first publication for works made for hire.³¹ To the extent that digitization is infringing, libraries and technology developers cannot afford to ignore the fact that these works may be subject to copyright because, even in the absence of actual harm or malicious intent, copyright owners may recover both statutory damages (up to \$150,000 per work infringed) and attorney’s fees.³²

The scale of the orphan works issue is potentially vast. One estimate finds that only 2.3% of books published in the United States between 1927 and 1946 are still in print.³³ Another reports that five out of every seven books Google scanned were not commercially available.³⁴ The Authors Guild estimates that approximately 75% of books in U.S. libraries are out-of-print and have ceased earning any income at all for their rights holders.³⁵ As the

28. Motion for Preliminary Settlement Approval, *supra* note 12, at 4.

29. *See generally* REPORT ON ORPHAN WORKS, *supra* note 4, at 23–29.

30. *Id.* at 26–29.

31. 17 U.S.C. § 302(b)–(c) (2010). As Pamela Samuelson notes:

Had copyright terms not been repeatedly extended by Congress, all books published before 1953 would now be in the public domain, as would most of the books published before 1978 insofar as their rights holders did not renew the copyright. Because of copyright term extensions, books first published in 1960 are, however, unlikely to be out of copyright until 2055.

Pamela Samuelson, *Google Book Search and the Future of Books in Cyberspace*, 94 MINN. L. REV. 1308, 1313 (2010); *see also* Christopher Sprigman, *Reform(aliz)ing Copyright Law*, 57 STAN. L. REV. 485 (2004); *see generally* REPORT ON ORPHAN WORKS, *supra* note 4, at 41–44.

32. 17 U.S.C. §§ 504(c) (statutory damages), 505 (attorney’s fees).

33. *See* Jason Schultz, *The Myth of the 1976 Copyright “Chaos” Theory*, LESSIG 2.0, <http://www.lessig.org/blog/archives/jasonfinal.pdf> (last visited June 14, 2012).

34. *See* Motoko Rich, *Google Hopes to Open a Trove of Little-Seen Books*, N.Y. TIMES, Jan. 5, 2009, at B1.

35. Motion for Preliminary Settlement Approval, *supra* note 12, at 27.

Copyright Office report on orphan works notes, this problem is particularly severe for institutions, such as libraries and museums, whose mission is to preserve and make available large archives of historical works.³⁶

To the extent that rights clearance is truly uneconomic, copyright is failing both orphan works owners and the public at large. Copyright exists to enable authors to set a price on access, not to frustrate access for its own sake. Library digitization's enormous potential (whether it be economic, educational, social, or democratic), and the copyright law's current failure in relation to orphan works, have led to many proposals for reform.³⁷

C. LIBRARY DIGITIZATION-SEARCH

The GLP version proposed under the ASA requires either judicial approval of the ASA (which will not be forthcoming) or legislative intervention. But what if Google were to scale back its ambitions to its initial proposal where unauthorized digitization was only incident to search? In the pure search scenario, the legality question of library digitization initiatives takes on a different complexion. Stated briefly, the argument favoring the legality of scanning, processing, and making fractional displays of books involved in GLP-search has three significant parts.

First, search results consisting of bibliographic information and relevance to a particular search query are facts not subject to copyright protection.³⁸ This is textbook copyright law in the United States and beyond serious dispute.³⁹

Second, the very brief snippets or quotations that Google displays in its search results are either (a) too brief, fragmented, and insubstantial to constitute a reproduction of an entire copyrighted work⁴⁰ or (b) used in a transformative manor to indicate the relevance of search results and not to substitute for the actual text, as such these snippets serve a different function

36. REPORT ON ORPHAN WORKS, *supra* note 4, at 37–38.

37. *See, e.g.*, Shawn Bentley Orphan Works Act of 2008, S. 2913, 110th Cong. (2008).

38. 17 U.S.C. § 102(b) (providing that “[i]n no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such a work”). *See also* Feist Publ'g, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 363–64 (1991) (noting that copyright distinguishes between facts and their expression).

39. *See infra* Section III.A for further discussion of the idea-expression distinction.

40. Even if one took the view that Google's actual three line snippets were too long, there must be *some* length of snippet—whether it be three lines, two lines, one line, or ten words—that would be non-infringing.

than the original work and are thus fair use.⁴¹ Either conclusion renders the search results displayed in GLP-search non-infringing. However, even if Google never showed a single book to anyone, the fact remains that it has been technically copying entire works to create its searchable database.

Third, copying entire expressive works for non-expressive (and otherwise non-infringing) purposes is itself fair use.⁴² Notice here that although orphan works may raise distinct issues in some contexts, the legitimacy of scanning and digitizing orphan works for library digitization-search is largely folded into the broader question of the scope of the copyright owner's rights in relation to non-expressive use. However, we should not lose sight of the importance of orphan works to the underlying policy debate. The intractable licensing problems that create orphan works mean extending the rights of copyright owners to include non-expressive use that would create a substantial market failure. Going forward, it is conceivable that publishers will get the rights they need from authors and agree to license these rights to those seeking to make non-expressive use of covered works, but, for the reasons canvassed above, the rights with respect to the majority of orphan works held in libraries will never be cleared. Put simply, if a court establishes a new right that gives copyright owners a veto over non-expressive use of their works, those rights will never be cleared for millions of orphan works.

III. NON-EXPRESSIVE USE

A. COPYRIGHT, BALANCE, AND THE DISTINCTION BETWEEN IDEAS AND EXPRESSION

As expressed in the U.S. Constitution, copyright's motivating purpose is "to promote the Progress of Science and useful Arts."⁴³ Copyright law in the United States does not exist primarily to recognize or validate the natural rights of authors vis-à-vis their creations. Instead, its purpose is to encourage the authors' creativity and to promote the creation and dissemination of

41. See *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007) (holding that "[e]ven making an exact copy of a work may be transformative so long as the copy serves a different function than the original work.") (citing *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818–19 (9th Cir. 2003)); see also *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 609–10 (2d Cir. 2006); *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1117–19 (D. Nev. 2006).

42. There are at least three search engine cases indicating as much. See *Perfect 10*, 508 F.3d at 1167–68; *Kelly*, 336 F.3d at 815; *Field*, 412 F. Supp. 2d at 1117–19. For a discussion of the fair use implications of non-expressive use generally, see Matthew Sag, *Copy-Reliant Technology*, *supra* note 7.

43. U.S. CONST. art. I, § 8, cl. 8.

works of authorship.⁴⁴ As the Supreme Court has noted on a number of occasions, the promotion of science and the useful arts requires a balance between “the interests of authors and inventors in the control and exploitation of their writings and discoveries on the one hand, and society’s competing interest in the free flow of ideas, information and commerce on the other hand.”⁴⁵ Where the law strikes that balance dictates what the public can copy and what authors can control. Just as importantly, it also mediates relationships between different generations of authors: initial authors and those who build upon their works.⁴⁶ Thus, while copyright aims to give authors an incentive to create and share their works, it also strives to provide subsequent authors with sufficient “breathing room” to make their own additive contributions.⁴⁷ The copyright system is predicated both on the existence of certain rights to protect authors from unfair competition, and on significant gaps in those rights that give others freedom to create and freedom to interact with existing works.

The distinction between ideas and expression is an important part of the balance of copyright law.⁴⁸ Copyright in an expressive work does not confer any exclusive rights in the facts, ideas, concepts, or discoveries contained in that work, regardless of the form in which the work describes, explains, or illustrates them.⁴⁹ This principle, often simply abbreviated to the “idea-

44. *Eldred v. Ashcroft*, 537 U.S. 186, 219 (2003).

45. *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 580 (1985) (quoting *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 429 (1984)).

46. See generally Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989 (1997) (discussing sequential innovation in copyright and patent law).

47. See, e.g., *Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 933 (2005). The court in *Sony* noted:

The fair use doctrine must strike a balance between the dual risks created by the copyright system: on the one hand, that depriving authors of their monopoly will reduce their incentive to create, and, on the other, that granting authors a complete monopoly will reduce the creative ability of others.

Sony, 464 U.S. at 479.

48. *Eldred*, 537 U.S. at 219 (stressing that the idea-expression distinction is one of copyright’s “built-in First Amendment accommodations” and that “[d]ue to this distinction, every idea, theory, and fact in a copyrighted work becomes instantly available for public exploitation at the moment of publication.”); *Harper & Row*, 471 U.S. at 556 (noting that the idea-expression distinction “strikes a definitional balance between the First Amendment and the Copyright Act by permitting free communication of facts while still protecting an author’s expression.”).

49. 17 U.S.C. § 102(b) (2010); *Harper & Row*, 471 U.S. at 547 (1985) (holding that “no author may copyright facts or ideas”).

expression distinction,” is longstanding at common law and was expressly incorporated into the 1976 revision of the Copyright Act.⁵⁰

At least since *Baker v. Selden* in 1879, courts have recognized that “there is a clear distinction between the book, as such, and the art that it is intended to illustrate.”⁵¹ The distinction holds even in those unusual cases where the work’s true value lies in the methods, systems, and ideas it discloses, rather than in the author’s expression of those concepts.⁵² In *Selden*, for example, the plaintiff had developed a novel and useful bookkeeping method, the practice that created value regardless of how and from what source a bookkeeper learned the method.⁵³ Nonetheless, the plaintiff’s copyright in his instructional material was limited to the expression of his useful methods and did not encompass those methods themselves.⁵⁴ Of course, in most cases, protecting the unique expression of an idea is sufficient to ensure that the author will be able to appropriate a return on her investment.⁵⁵

Copyright law also distinguishes between facts and the expression of facts, providing no protection for the former and only limited protection for the latter.⁵⁶ In *Feist Publications, Inc. v. Rural Telephone Service Co.*, the Supreme Court ruled that copying listings from a telephone directory did not infringe the copyright in that directory because the information itself was not copyrightable.⁵⁷ As the Court explained, facts—whether they are telephone numbers and addresses or the details of historical occurrences—are not

50. 17 U.S.C. § 102(b) provides: “In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such a work.”

51. *Baker v. Selden*, 101 U.S. 99, 102 (1879). “Art” and “illustrate” are not meant in the aesthetic sense in this context.

52. *Id.*

53. *Id.* at 99–100. *Selden*’s system may well have been patentable under today’s standards. *See State Street Bank & Trust Co. v. Signature Fin. Group, Inc.*, 149 F.3d 1368, 1373 (Fed. Cir. 1998) (holding that a patent on a data processing system is valid). *But see Lab. Corp. of Am. Holdings v. Metabolite Labs., Inc.*, 548 U.S. 124, 136 (2006) (per curiam) (Breyer, J., dissenting) (noting that the Supreme Court has never endorsed the Federal Circuit’s “useful, concrete, and tangible result” test for patentable processes). *See generally* Pamela Samuelson, *Why Copyright Law Excludes Systems and Processes from the Scope of Its Protection*, 85 TEX. L. REV. 1921, 1924 (2007) (arguing that thin copyright protection for computer programs is especially appropriate given the availability of patent protection for program innovations).

54. *Baker*, 101 U.S. at 103–04.

55. *Cf.* WILLIAM M. LANDES & RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY* 91–108 (2003).

56. *See Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 349–50 (1991) (holding that facts are not copyrightable and that the copyright in a factual compilation is thin).

57. *Id.* at 362–63.

“original” to the author.⁵⁸ The author’s copyright, therefore, did not cover the facts themselves.⁵⁹ The *Feist* Court further held that the expression of those facts was not protectable, because the selection and alphabetical arrangement of those facts in the telephone directory was “so mechanical or routine as to require no creativity whatsoever.”⁶⁰ The rule in *Feist* even extends to “false facts.”⁶¹

Through the idea-expression distinction, copyright law protects the expressive elements of the author’s work while guaranteeing subsequent authors the necessary breathing space to make their own contributions by adding to, reusing, or reinterpreting the facts and ideas embodied in the original work. Subsequent authors may not compete with the copyright owner by offering her original expression to the public as a substitute for the copyright owner’s work, but they are free to compete with their own expression of the same facts, concepts, and ideas. Thus, the idea-expression distinction is a central element of the balance between the interests of authors in preventing the exploitation of their writings and society’s competing interest in the free flow of ideas, information, and commerce.⁶²

Demarcating the precise boundary between ideas and expression is no easy task. The famous 1930 case of *Nichols v. Universal Pictures Corp.* dealt with a play about lovers from different religious backgrounds and a motion picture with the same motif.⁶³ The playwright, whose work came first, alleged

58. *Id.* at 348 (“[C]opyright protection may extend only to those components of a work that are original to the author.”).

59. *Harper & Row*, 471 U.S. at 556 (“No author may copyright his ideas or the facts he narrates.”).

60. *See Feist*, 499 U.S. at 362 (holding that the selection, coordination, and arrangement of Rural’s white pages did not satisfy the minimum constitutional standards for copyright protection); *see also* *Matthew Bender & Co. v. West Publ’g Co.*, 158 F.3d 674, 676 (2d Cir. 1998) (holding that West’s factual enhancements to judicial opinions could be reasonably viewed as obvious, typical, and lacking even minimal creativity).

61. False facts are denied protection under a theory of “copyright estoppel.” *Skinder-Strauss Assocs. v. Mass. Continuing Legal Educ., Inc.*, 914 F. Supp. 665, 675–76 (D. Mass. 1995); *Houts v. Universal City Studios, Inc.*, 603 F. Supp. 26, 28 (C.D. Cal. 1984) (“once a plaintiff’s work has been held out to the public as factual the author-plaintiff cannot then claim that the book is, in actuality, fiction and thus entitled to the higher protection allowed to fictional works.”). Some courts have been willing to grant de facto database protection to individual facts brought into being as a result of creative choices, such as bluebook valuations, and price guides to rare coins. *See, e.g., CDN, Inc. v. Kapes*, 197 F.3d 1256 (9th Cir. 1999).

62. *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 429 (1984); *see also Warner Bros. v. Am. Broad. Cos.*, 720 F.2d 231, 240 (2d Cir. 1983) (describing the idea-expression distinction as “an effort to enable courts to adjust the tension between these competing effects of copyright protection”).

63. *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 120–21 (2d Cir. 1930).

that the movie infringed his rights.⁶⁴ Ruling for the defendant, Judge Learned Hand observed that although copyright must extend beyond the exact literal text of a work, similarities between two works at a high level of generality cannot violate the author's rights because a playwright can not "prevent the use of his 'ideas,' . . . apart from their expression"⁶⁵ Having described the idea-expression distinction, the learned judge immediately observed that "[n]obody has ever been able to fix that boundary, and nobody ever can."⁶⁶ Although the precise point of departure between protectable expression on the one hand and unprotectable fact and ideas on the other may be elusive,⁶⁷ unstable⁶⁸ and somewhat subjective,⁶⁹ no one doubts that it exists.⁷⁰

The distinction between expressive and non-expressive *works* is already well recognized in copyright law as the gatekeeper to copyright protection—novels are protected by copyright, telephone books and other uncreative compilations of data are not.⁷¹ The position of this Article is that the same distinction should generally be made in relation to potential *acts* of infringement. Preserving the functional force of the idea-expression distinction in the digital context requires courts to conclude that copying for purely non-expressive purposes, such as the automated data extraction, should not be regarded as infringing. As this Article will explain in Section III.C, *infra*, courts are already tacitly implementing the principle of non-expressive use in the case law. The principle, however, needs to be brought to the surface.

64. *Id.*

65. *Id.* at 121.

66. *Id.* See also *Peter Pan Fabrics, Inc. v. Martin Weiner Corp.*, 274 F.2d 487, 489 (2d Cir. 1960) ("Obviously, no principle can be stated as to when an imitator has gone beyond copying the 'idea,' and has borrowed its 'expression.' Decisions must therefore inevitably be ad hoc.").

67. Professor Chafee's proposed "pattern" test for determining the line between an idea and its expression is as good as any, but it essentially reframes the question rather than answering it. Zechariah Chafee, *Reflections on the Law of Copyright: I*, 45 COLUM. L. REV. 503, 513–14 (1945).

68. See, e.g., Peter Jaszi, *The Metamorphoses Of "Authorship"*, 41 DUKE L.J. 455, 465 (1991).

69. See Amy B. Cohen, *Copyright Law and the Myth of Objectivity: The Idea-Expression Dichotomy and the Inevitability of Artistic Value Judgments*, 66 IND. L.J. 175, 228 (1990) (reviewing the application of the idea-expression distinction in case law and concluding that where the line is drawn "reflects the judge's view of the artistic value of the works at issue based on what the judge knows about and values in literary works on that subject.").

70. But see Edward C. Wilde, *Replacing the Idea/Expression Metaphor With a Market-Based Analysis in Copyright Infringement Actions*, 16 WHITTIER L. REV. 793 (1995).

71. *Feist Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340 (1991).

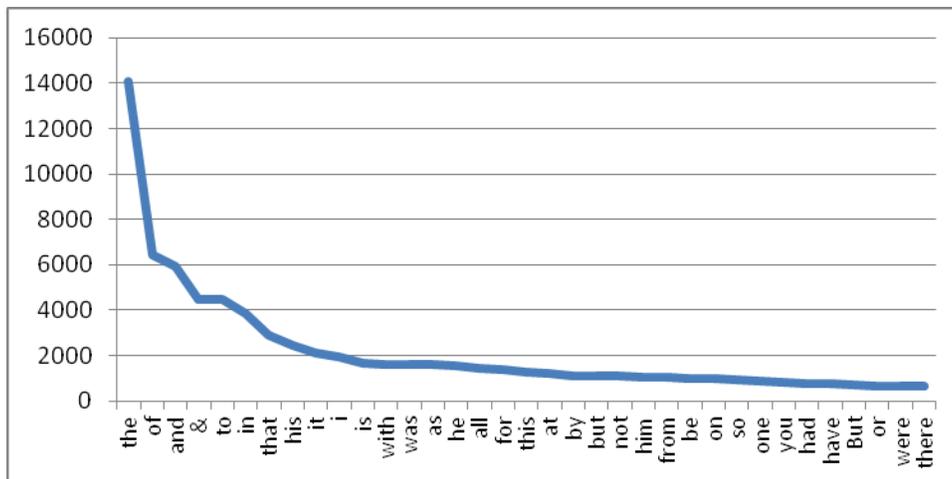
B. NON-EXPRESSIVE USE

1. *Coming to Grips with the Concept of Non-Expressive Use*

To understand what non-expressive use means, consider the following thought experiment.

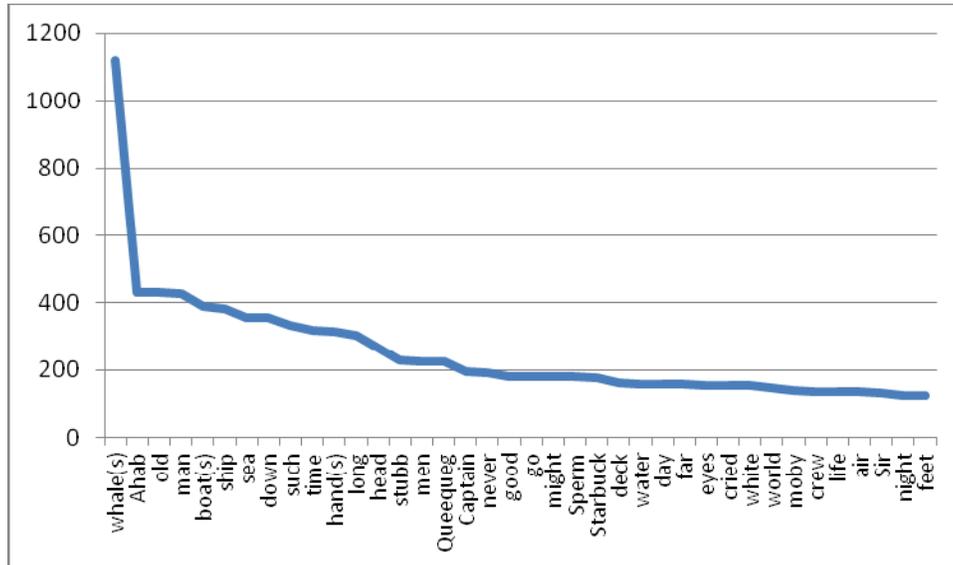
Brian has a perfect memory for the written word: he can recite every book he has ever read perfectly from start to finish. He can, if pushed, write out frequency tables that report the number of times any given word or punctuation mark appears in any work.⁷² Brian might, for example, produce the following word frequency graphs for Herman Melville's *Moby Dick*. Figure 1 illustrates the frequency of common English words in *Moby Dick* based on a list of words that is not sensitive to context such as “the,” “of,” “and,” “have,” etc. Figure 2 illustrates word frequencies in *Moby Dick* once the words in Figure 1 have been excluded.

Figure 1: Frequency of Common English Words in *Moby Dick*⁷³



72. ALEKSANDR ROMANOVICH LURIA, *THE MIND OF A MNEMONIST: A LITTLE BOOK ABOUT A VAST MEMORY* (1987) (an account of a Russian man with a limitless memory).

73. HERMAN MELVILLE, *MOBY DICK; OR, THE WHALE*, available at <http://www.princeton.edu/~batke/moby/moby.html>. Word frequency obtained using Wordle.net java applet. WORDLE, <http://www.wordle.net/> (java applet for obtaining word frequency). Words selected by the author.

Figure 2: Frequency of Selected Words in Moby Dick⁷⁴

The same information can be represented in a more whimsical visual style using a word cloud as follows in Figure 3 and Figure 4.

74. HERMAN MELVILLE, MOBY DICK; OR, THE WHALE, *available at* <http://www.princeton.edu/~batke/moby/moby.html>. Word frequency obtained using Wordle.net java applet. WORDLE, <http://www.wordle.net/> Words selected by the author.

copyright.⁷⁸ But would the frequency table infringe the author's copyright? The frequency table itself is metadata, data about the work that is entirely independent of the expressive value of the work. True enough, the data relies on the underlying work, but it has no similarity to the work in terms of plot, structure, character (other than the names of characters) or theme. This data, by itself, does not infringe the copyright owner's rights.

Is there a point at which an analytical work explains so much of the content of its expressive subject that the author's rights have been infringed? Perhaps. In *Castle Rock Entertainment v. Carol Publishing Group, Inc.*,⁷⁹ the Second Circuit held that a quiz book based on the characters and events of the popular television series *Seinfeld* violated the show's copyright. The court acknowledged that the substantially similar standard depends on "the copying of expression, rather than ideas" and that the quiz reproduced none of the plot, sequence, pace, or setting of the show.⁸⁰ The defendant's quiz focused on "facts" internal to the *Seinfeld* universe, such as the reason that Kramer enjoys going to the airport (because he is hypnotized by the baggage carousels) or what it was that Jerry placed on Elaine's leg during a piano recital (a Pez dispenser), and not facts about the show.⁸¹ The court of appeals took the view that "[b]ecause these characters and events spring from the imagination of *Seinfeld's* authors, the [quiz] plainly copies copyrightable, creative expression."⁸² Of course, the court cannot really mean that any work that refers to the characters and events in a creative work is infringing. Furthermore, there are volumes of guide-books and analytical works that do not interfere with the copyright owner's exclusive rights, and it is well established that "ownership of copyright does not confer a legal right to control public evaluation of the copyrighted work."⁸³ The real problem with the defendant's quiz in *Castle Rock Entertainment* was that it sought to "repackage *Seinfeld* to entertain *Seinfeld* viewers" and that the quiz itself was in no way analytical.⁸⁴ If the *Seinfeld* quiz infringed the copyright owner's rights at all, it was because it essentially recast the series' copyrightable characters

78. 17 U.S.C. § 106(1) (reproduction), (3) (distribution).

79. 150 F.3d 132 (2d Cir. 1998).

80. *Id.* at 138.

81. *Id.* at 139.

82. *Id.*

83. *Ty, Inc. v. Publ'ns Int'l. Ltd.*, 292 F.3d 512, 521 (7th Cir. 2002).

84. *Castle Rock Entertainment*, 150 F.3d at 140–43.

into a new format, much the same as if the defendant had made miniature dolls of the show's characters.⁸⁵

The recent *Harry Potter Lexicon* case is also on point.⁸⁶ In *Warner Brothers Entertainment Inc. v. RDR Books*, the court found that a guidebook to the famed *Harry Potter* series violated the author's copyright.⁸⁷ The court found that the Lexicon was substantially comprised of direct quotations (often without quotation marks) and close paraphrases of vivid passages in the *Harry Potter* books.⁸⁸ Like the *Seinfeld* quiz, the Lexicon related "fictional facts" the author, J.K. Rowling, had created. In line with *Castle Rock*, the court concluded "such invented facts constitute creative expression protected by copyright because characters and events spring from the imagination of the original authors."⁸⁹ One interpretation of the court's opinion in the *Harry Potter Lexicon* case is that if the guidebook had not borrowed so extensively from the original author's expression, it would not have been found to infringe.⁹⁰ The Lexicon's purpose was to "give the reader a ready understanding of individual elements in the elaborate world of *Harry Potter* that appear in voluminous and diverse sources."⁹¹ The district court in the *Harry Potter Lexicon* case held that the Lexicon did not infringe the copyright owner's right to make derivative works because it no longer represented the original work of authorship and did not fall under any example of derivative works listed in the statute.⁹² The court followed the Seventh Circuit's holding that a collector's guide to stuffed toys is not a derivative work because "guides don't recast, transform, or adapt the things to which they are guides."⁹³ If the Lexicon had been drafted with more care, it need not have infringed the copyright owner's rights.

85. See, e.g., *Hasbro Bradley, Inc. v. Sparkle Toys, Inc.*, 780 F.2d 189 (2d Cir. 1985) (upholding copyrightability of "Transformer" changeable robotic action figures as sculptural works).

86. *Warner Bros. Entm't Inc. v. RDR Books*, 575 F. Supp. 2d 513 (S.D.N.Y. 2008).

87. *Id.*

88. *Id.* at 527 ("the Lexicon indeed contains at least a troubling amount of direct quotation or close paraphrasing of Rowling's original language"); *id.* at 530 ("The Lexicon's close paraphrasing is not limited to the seven *Harry Potter* novels, but can be found in entries drawn from the companion books as well."); *id.* at 531 ("Instances of such verbatim copying or close paraphrasing of language in the *Harry Potter* works occur throughout the Lexicon.").

89. *Id.* at 536 (citations and quotations omitted).

90. The decision could be clearer as to the relationship between findings of fact and legal conclusions.

91. *Id.* at 539.

92. *Warner Bros.*, 585 F. Supp. 2d at 539 ("Under these circumstances, and because the Lexicon does not fall under any example of derivative works listed in the statute, Plaintiffs have failed to show that the Lexicon is a derivative work.").

93. *Ty, Inc. v. Publ'ns Int'l. Ltd.*, 292 F.3d 512, 520 (7th Cir. 2002).

The automated data analysis of text that this Article addresses is a far cry from the fragmented expression copying in the *Harry Potter Lexicon* case and other similar “fictional facts” cases. Copyright does not protect individual words, even in the rare instances where they are in fact a creation of the author.⁹⁴ For example, an author such as J.K. Rowling can be said to originate the following twenty-word string of text: “[g]oblin-made armour does not require cleaning, simple girl. Goblins’ silver repels mundane dirt, imbibing only that which strengthens it.”⁹⁵ But none of these individual words originates with Rowling. The corresponding entry in the Lexicon reads “[a]ccording to Phineas Nigellus, goblin-made armor does not require cleaning, because goblins’ silver repels mundane dirt, imbibing only that which strengthens it, such as basilisk venom.”⁹⁶ Moreover, the observation that no word other than “goblin” is repeated in either sentence originates, not with Rowling, but with the author of this Article.⁹⁷ Likewise, if some anti-plagiarism software were to identify a high level of similarity between the two quotes—as it surely would—that data could not be said to originate with either the author of *Harry Potter* or the author of the Lexicon. It is a fact about the works and is in no sense a reproduction of either work or a substantial part of the original expression therein. In summary, metadata of the sort described here infringes only as much as a landscape painting inspired by a novel, or a musical composition inspired by a film would—i.e., not at all.⁹⁸

Returning to our thought experiment, would Brian infringe the copyright owners rights by simply memorizing *Moby Dick* as part of the process of making the table? If Brian is a human being, it seems absurd to suggest that the perfect storage of information in his brain violates the copyright owner’s exclusive right to “reproduce the work in copies . . .” under § 106(1) of the Copyright Act. Even if scientists told us that Brian’s brain stored and could recall the information with perfect accuracy,⁹⁹ it is inconceivable that human

94. The Copyright Office has a long-standing rule that “words and short phrases such as names, titles, and slogans” are not copyrightable. 37 C.F.R. § 202.1(a) (2004). See Justin Hughes, *Size Matters (Or Should) In Copyright Law*, 74 FORDHAM L. REV. 575 (2005).

95. *Warner Bros.*, 575 F. Supp. 2d at 527 (quoting J.K. ROWLING, HARRY POTTER AND THE DEATHLY HALLOWS 303 (2007)).

96. *Id.*

97. Admittedly, this is not a profound observation.

98. See Robert Kastenmeier, Copyright Law Revision, H.R. Rep. No. 94-1476 (2d sess. 1976) (noting a programmatic musical composition inspired by a novel).

99. Looking closely at the definition of copies in § 101 of the Act it is not immediately clear that the human brain cannot be a copy. To amount to a copy under the Act, the medium storage must simply be a “material object . . . in which a work is fixed . . . and from which the work can be perceived, reproduced, or otherwise communicated, either directly or

thought or human memory could be a form of copyright infringement.¹⁰⁰ Now suppose that Brian is a computer; should the answer really be any different?

2. *Examples of the Non-Expressive Use of Expressive Works*

Ordinarily, the direct or indirect purpose of reproducing an expressive work relates to human appreciation of the expressive qualities of that work. We might, for example, download a film to watch it, or photocopy a magazine article to read it. The examples that follow illustrate a very different kind of motivation for copying text: reproduction as part of a process of data analysis that does not enable human enjoyment, appreciation, or comprehension of the text. These examples demonstrate the utility of automated non-expressive uses. They also demonstrate that such uses are no threat to the interests of copyright owners. This Section begins with two of the more obvious examples unrelated to library digitization—Internet search engines and plagiarism detection software—before turning to the role of non-expressive use in library digitization.

a) Internet Search Engines¹⁰¹

Internet search engines provide the most obvious example of the importance of the non-expressive use of copyrighted works. Internet search engines are a form of copy-reliant technology in that they require the routine and indiscriminate copying of html web pages.¹⁰² Search engines use automated software agents that continuously “crawl” across the Internet copying web pages. These copies form the raw data underpinning these

with the aid of a machine or device.” 17 U.S.C. § 101 (2010). Most people’s brains do not store information with the stability and fidelity required to meet this definition, but what of those that do? One answer is to posit, as David Nimmer has in relation to tattoos, that a human is not a “material object,” and while this may be a sound policy-based exclusion, it does not supply its own rationalization. *See* Declaration of David Nimmer, *Whitmill v. Warner Bros. Entm’t*, No. 4:11CV752 (E.D. Mo. May 20, 2011) (declaration in support of the defendant in copyright litigation regarding the use of Mike Tyson’s facial tattoo in the motion picture *THE HANGOVER II* (Warner Brothers 2011)).

100. Likewise in patent law, “[a] principle, in the abstract, is a fundamental truth; an original cause; a motive; these cannot be patented, as no one can claim in either of them an exclusive right.” *Le Roy v. Tatham*, 55 U.S. (14 How.) 156, 175 (1853); *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 132 S. Ct. 1289 (2012) (noting that natural phenomena, mental processes, and abstract intellectual concepts are not patentable). *See generally* Kevin Emerson Collins, *Propertizing Thought*, 60 SMU L. REV. 317 (2007).

101. *See* Sag, *Copy-Reliant Technology*, *supra* note 7 for a more detailed account the operation of Internet search engines and plagiarism detection software.

102. Sergey Brin & Lawrence Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, 30 COMPUTER NETWORKS AND ISDN SYS. 107 (1998), available at <http://infolab.stanford.edu/~backrub/google.html>.

search engines, which are subsequently analyzed and cataloged. As part of this process, search engines both copy and index each web page they find. The search engine directs the user to particular websites based on the relationship of her search term to the index of pages maintained by the search engine provider.¹⁰³ The search engine's use is non-expressive because the software copies expressive works in order to apply certain mathematical functions to their contents, not to comprehend or enjoy copyrighted works in the way that humans do. Of course, at the end of the day, search engines are mostly useful because they lead people to particular websites. But the search engine itself does not copy the website for the end user. Instead, this process is performed separately by the user's browser at the direction of the user.¹⁰⁴

b) Plagiarism Detection Software

Plagiarism detection software is another illustration of the copying of expressive works for non-expressive ends. In the educational context, automated plagiarism services rely on access to entire copies of student term papers and any works from which a student might have copied them, yet the services do not necessarily display any of the copyrighted content they process to the end users.¹⁰⁵ The software works by comparing strings of text in new works to strings of text in existing works.¹⁰⁶ The similarities between two works can be assessed by looking for common strings of words. However, there are also various algorithms that can be applied to a document to create a digital fingerprint that captures other characteristics of the work. These digital fingerprints allow a document to be characterized by its structure, vocabulary, and content. Furthermore, they are essentially abstractions of the original documents and allow for faster comparisons that will not be as easily deceived by minor text alterations.¹⁰⁷ If the software finds

103. See, e.g., U.S. Patent No. 6,285,999 (filed Jan. 9, 1998) ("Method for Node Ranking in a Linked Database").

104. *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1161 (2007).

105. See Sag, *Copy-Reliant Technology*, *supra* note 7 for a more detailed account the operation of Internet search engines and plagiarism detection software.

106. See Amy Argetsinger, *Technology Snares Cheaters at U-Va.; Physics Professor's Computer Search Triggers Investigation of 122 Students*, WASH. POST, May 9, 2001, at A1.

107. See, e.g., Khair Eddin M. Sabri & Jubair J. Al-Ja'afar, *The JK System to Detect Plagiarism*, 6(2) J. COMPUTER SCI. & TECH. 66 (2006). The Turnitin software at issue in *A.V. ex rel. Vanderbye v. iParadigms, LLC*, 562 F.3d 630 (4th Cir. 2009) used statistical techniques originally designed to analyze brain waves to compare the fingerprints of student papers to more than a billion documents that have been fingerprinted in a similar fashion. See *Plagiarise. Let No One Else's Work Evade Your Eyes*, THE ECONOMIST, Mar. 14, 2002, available at <http://www.economist.com/node/1033832>.

a match, it indicates as much. By itself, the report that a new work is similar to another work already in the database in no way reproduces or communicates the expressive qualities of either work.¹⁰⁸

c) Non-Expressive Use and Library Digitization

Library digitization raises many novel issues, but one should not lose sight of the fact that some of the relevant issues are not at all new. The fundamental issue with respect to the legality of copying to build a search engine is the same for web pages as it is for library books. In point of fact, there are some interesting differences. To start, library digitization also raises interesting questions about the scope of the § 108 library privilege.¹⁰⁹ Non-profit libraries that undertake digitization initiatives might have additional arguments to make with respect to fair use. Likewise, the automated copying of html pages may also be protected by an implied license in many cases.¹¹⁰ But these are distractions; the key question remains whether automated and systematic copying of text to enable a search engine (but not a display engine) or other data-processing function violates the rights of the copyright owner.

In addition to book searches, there are many non-expressive uses for library digitization. Researchers could use a digitized collective (referred to in the trade as the “corpus”) to test and refine search algorithms more generally.¹¹¹ Other researchers could use the resulting data field to improve automated translation software and to develop and test theories in linguistics. Some of the most interesting illustrations of the kind of non-expressive use that library digitization enables relate to the meta-analysis of literature.

In the world of books, a non-expressive use is any use that, while it may literally involve reproduction of the work, does not involve any human

108. Of course, in practice most plagiarism software is also programmed to display the source file from which the work being scrutinized was allegedly copied. This optional feature is an expressive use, although it is almost certainly protected by fair use because the purpose of the display is to provide evidence of a claim of cheating. *A.V. ex rel. Vanderhye*, 562 F.3d at 641–42 (finding that the defendant’s use of the works as part of a digitized database from which to compare the similarity of typewritten characters used in other student works was unrelated to any creative component of the work).

109. See, e.g., Peter B. Hirtle, *Digital Access to Archival Works: Could 108(b) Be the Solution?*, COPYRIGHT & FAIR USE: STANFORD UNIV. LIBRARIES (Sept. 24, 2006), http://fairuse.stanford.edu/commentary_and_analysis/2006_08_hirtle.html.

110. *Field v. Google Inc.*, 412 F. Supp. 2d 1106, 1115 (D. Nev. 2006).

111. See ASA, *supra* note 23, § 1.93 (defining non-consumptive use to include Image Analysis and Text Extraction, Textual Analysis and Information Extraction, Linguistic Analysis, Automated Translation, and Indexing and Search (research on different techniques for indexing and search of textual content)).

reading the digitized copy of the book. If the data extracted does not allow for the work to be reconstructed, there is no substitution of expressive value. Extracting factual information about a work in terms of its linguistic structure or the frequency of the occurrence of certain words, phrases, or grammatical features is a non-expressive use.¹¹²

To start with a simple example, merely reporting the fact that the word “whale” or “whales” appears 1,119 times in Herman Melville’s *Moby Dick* does not infringe any copyright in the book because this information about the work is entirely independent of the expressive value of the work.¹¹³ There is no copyright in such basic information as the names of characters in a novel or a list of places they have been.¹¹⁴ Nor is copyright infringed by the simple observation that Melville writes a great deal about whales, old men, the sea, boats, water, and ships. To preserve the force of the idea-expression distinction in the age of reading machines, one must recognize that copyright law does not prevent the automated extraction of such features by machine applications, even if those machines reproduce the text as a step in the analytical process. In this context, so long as the output is non-infringing, the machine is non-infringing.

Consider, for example, Franco Moretti’s fascinating map of protagonists in Parisian Novels and the objects of their desire.¹¹⁵ Aggregating information across many books allows us to see not only that the heroes of this particular genre are clustered in the Latin Quarter, but also that they are invariably separated from their heart’s true desire by the River Seine and distributed in a convex arc as if held from the Latin Quarter by a constant unseen force. Moretti and a team of graduate students constructed this map by hand, but there is no obvious reason why a similar process on a grander scale could not be automated.

112. The ASA uses the awkward term “Non-Consumptive Research” to express the same concept. The ASA defines Non-Consumptive Research as “research in which computational analysis is performed on one or more Books, but not research in which a researcher reads or displays substantial portions of a Book to understand the intellectual content presented within the Book.” *Id.*

113. *See supra* Figure 2.

114. For a literary character to be protected as such by copyright it must, at a minimum, be distinctively delineated such that it represents a specific incarnation and not a general archetype. *Warner Bros. Pictures, Inc. v. Columbia Broadcasting Systems, Inc.*, 216 F.2d 945, 950 (9th Cir. 1954) sets a higher standard, that the “the character really constitutes the story being told” and is not merely a “chess man in the game of telling the story.”

115. FRANCO MORETTI, *GRAPHS, MAPS, TREES* 55 (2005).

Literature scholars have traditionally focused on a close reading of canonical texts as the core of their discipline.¹¹⁶ Even those who venture further afield do not travel that far. For example, literary historian Ian Watt's seminal 1957 work on the origins of the novel¹¹⁷ is undoubtedly a brilliant synthesis of modern literature, and yet his entire scope of analysis is confined to three authors.¹¹⁸ Three! Close reading of the literary cannon or of a few dozen exemplar works in a particular literary genre is no doubt an impressive skill, but the limits of this analysis are striking.¹¹⁹ Empirically, the cannon itself cannot tell you anything about the characteristics that propel works into the cannon in the first place. To make that judgment, you must compare the cannon to the mass of other works all vying for that status but now largely forgotten.

Data-mining and macro-analysis of literature offers broad possibilities. Computer assisted text-analysis not only stores, searches, and retrieves text efficiently, it also automates the process of measuring and classifying natural-language documents to identify patterns that may be associated with author, subject, and genre or type.¹²⁰ Macro-analysis does not replace reading altogether, but it offers scholars a way to empirically test intuitions that are in fact quantitative or comparative in nature.¹²¹ To take a rudimentary example, the notion that female characters are underrepresented in a particular period may be intuited from a small selection of prominent works. As such, macro-analysis of all the books from that period would allow that intuition to be tested empirically and potentially confirmed or falsified.¹²² In his forthcoming book, *Macroanalysis: Methods for Digital Literary History*, Matthew Jockers uses various empirical techniques to identify the dominant themes in two of the most famous American novels of the nineteenth century—*The Last of the Mohicans* (1826) and *Moby Dick* (1851)—and contrast them against the nineteenth century corpus as a whole. Jockers does not read all 10,000 novels of the era,¹²³ but instead undertakes this investigation using word frequency

116. See generally MATTHEW JOCKERS, *MACROANALYSIS: DIGITAL METHODS AND LITERARY HISTORY* (forthcoming 2013).

117. IAN WATT, *THE RISE OF THE NOVEL* (1957).

118. Daniel Defoe, Samuel Richardson and Henry Fielding. See *id.* at 7.

119. There is an obvious parallel here with the rationale for conducting empirical legal studies. See, e.g., Matthew Sag, Tonja Jacobi & Maxim Sytch, *Ideology and Exceptionalism in Intellectual Property: An Empirical Study*, 97 CALIF. L. REV. 801 (2009).

120. See Geoffrey Rockwell, *Why Bother With Computer-Assisted Text Analysis? A Short Answer*, TEXT ANALYSIS DEVELOPERS ALLIANCE (Apr. 30, 2005), <http://tada.mcmaster.ca/Main/WhatTA>.

121. *Id.*

122. *Id.*; JOCKERS, *supra* note 116.

123. 10,000 is a very rough guess.

analysis and computer generated topic modeling that identifies patterns based on the frequency with which words are combined. Jockers is an English professor, but he borrows techniques developed in computational linguistics and natural language processing to take account of grammatical structure and idiomatic usage in this analysis. No doubt, this is just the beginning of an exciting new field. The question for lawyers, judges, and legal academics is whether this type of analysis must be limited to public domain works and those licensed by publishers.

C. THE SCOPE OF COPYRIGHT WITH RESPECT TO THE NON-EXPRESSIVE USE OF EXPRESSIVE WORKS

The prescription in this Article, that copyright law should not stand in the way of the automated reproduction of text for non-expressive purposes, rests on the view that, in general, the copyright owner's exclusive rights are limited to the right to communicate the expressive aspects of her work to the public. To put it another way, copyright typically only concerns itself with the threat of expressive substitution. As already noted, the idea-expression distinction itself establishes that the copyright owner cannot prevent an ordinary reader from extracting and reproducing the facts or ideas embodied in her work. But the principle goes much deeper than this.

Copyright consists of a bundle of discrete exclusive rights, such as the reproduction right, the derivative right, and the public performance and display rights.¹²⁴ These rights are defined, articulated, and limited by a number of initially judge-made doctrines, such as the idea-expression distinction, the threshold of substantial similarity, and the fair use doctrine.¹²⁵ In my earlier work, I have explained in some detail that these doctrines typically limit copyright protection to the expressive aspects of original works of authorship in a way that confirms the place of public communication at the heart of copyright.¹²⁶ This Article will expand and clarify just a few of these arguments.

1. *Substantial Similarity*

The tests courts apply to determine the threshold of infringement—i.e., when some copying is *too much* copying—strongly suggest that the statutory rights of the author are limited to the communication of original expression

124. 17 U.S.C. § 106(1)–(6) (2010).

125. The Copyright Act of 1976 also reflects the idea-expression distinction and the fair use doctrine. *See* 17 U.S.C. §§ 102(b), 107. But these doctrines remain essentially common law features of the copyright system.

126. *Sag, Copy-Reliant Technology*, *supra* note 7.

to the public. The copyright owner's exclusive right to "reproduce the copyrighted work in copies" extends to both exact and inexact reproductions.¹²⁷ In both cases, however, the Copyright Act leaves the threshold of reproduction undefined. In cases of nonliteral infringement—where the accused work is not an exact copy of the copyright owner's work—courts assess whether the allegedly infringing work possesses a substantial similarity to the copyrighted work.¹²⁸

Courts often define the threshold of substantial similarity from the perspective of the ordinary observer.¹²⁹ Infringement is defined in reference to the perspective of the consuming public because the copyright owner's "legally protected interest is not, as such, his reputation . . . but his interest in the potential financial returns from his [work] which derive from the lay public's approbation of his efforts."¹³⁰ Thus, the determination of whether work "B" borrowed too much from work "A" hinges upon how the public would regard the similarities between the works. But this is not the end of the analysis. Even when two works are similar taken as a whole, any similarities based on overlapping ideas or expressions that were not the plaintiff's to begin with "are by definition unprotected . . ." ¹³¹ A plaintiff in a copyright case "must show that defendants' works are substantially similar to elements of plaintiff's work that are *copyrightable* or protected by the copyright."¹³²

In cases of fragmented literal similarity, courts determine whether the copying amounts to infringement "by considering the qualitative and quantitative significance of the copied portion in relation to the plaintiff's work as a whole."¹³³ This focus on the qualitative and quantitative significance of the copied portion in the plaintiff's work is consistent with

127. 17 U.S.C. § 106(1); *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 121 (2d Cir. 1930) ("[T]he question is whether the part so taken is substantial" (citing *Marks v. Feist*, 290 F. 959, 960 (2d Cir. 1923))) (internal quotation marks omitted).

128. See *Tufenkian Imp./Exp. Ventures, Inc. v. Einstein Moomjy, Inc.*, 338 F.3d 127, 131 (2d Cir. 2003); *Laureyssens v. Idea Grp., Inc.*, 964 F.2d 131, 140 (2d Cir. 1992).

129. This is especially true in the Second Circuit. See *Shine v. Childs*, 382 F. Supp. 2d 602, 614 (S.D.N.Y. 2005) (summarizing authorities). For a survey of other approaches, see MELVILLE B. NIMMER & DAVID NIMMER, *NIMMER ON COPYRIGHT* § 13.03 (2012).

130. *Arnstein v. Porter*, 154 F.2d 464, 473 (2d Cir. 1946) (footnotes omitted); see also *Warner Bros. v. Am. Broad. Cos.*, 720 F.2d 231, 240 (2d Cir. 1983).

131. NIMMER *supra* note 129, § 13.03[2].

132. *Whitehead v. Paramount Pictures Corp.*, 53 F. Supp. 2d 38, 46 (D.D.C. 1999) (emphasis in original) (citing NIMMER, *supra* note 129, § 13.03[2]).

133. *Newton v. Diamond*, 388 F.3d 1189, 1195 (citing *Worth v. Selchow & Righter Co.*, 827 F.2d 569, 570 n.1 (9th Cir. 1987)); see also *Jarvis v. A&M Records*, 827 F. Supp. 282, 289–90 (D.N.J. 1993); NIMMER, *supra* note 129, § 13.03[A][2][a].

the prohibition against expressive substitution. Even where some of the copyright owner's original expression has been copied directly, such copying does not rise to the level of infringement unless the expression was significant, in either quantity or quality, in the author's original work.¹³⁴ Just as copyright law does not prevent the copying of facts and ideas, it also does not prevent the copying of trivial expressive elements from an existing work, because to do so does not unfairly compete with the copyright owner.¹³⁵ In other words, trivial copying of expressive elements is not copyright infringement because it does not interfere with the copyright owner's exclusive right to communicate her work to the public.

In summary, the very mechanics of assessing whether the threshold of substantial similarity has been met provide further evidence that copyright primarily protects the author against expressive substitution.

2. *Intermediate Copying*

For those in Hollywood, facing dubious claims of copyright infringement is a recognized cost of doing business.¹³⁶ Presumably, some of these claims are opportunistic, while others are the product of self-delusion. The attraction of a substantial payday combined with passing similarities based on title, theme, or subject matter can be enough to trigger a suit. What is significant for the purposes of this Article is that when confronted with motions for summary judgment based on an objective lack of similarity between their own work and that of the defendant, plaintiffs in a number of cases have turned to allegations of intermediate copying.¹³⁷ Typically, plaintiffs in this situation will urge the courts to allow scrutiny of every single

134. *Newton*, 388 F.3d at 1195 (9th Cir. 2004). The court noted:

Fragmented literal similarity exists where the defendant copies a portion of the plaintiff's work exactly or nearly exactly, without appropriating the work's overall essence or structure. Because the degree of similarity is high in such cases, the dispositive question is whether the copying goes to trivial or substantial elements. Substantiality is measured by considering the qualitative and quantitative significance of the copied portion in relation to the plaintiff's work as a whole.

Id. (internal citations omitted).

135. *Id.* at 1193, 1195 ("The principle that trivial copying does not constitute actionable infringement has long been a part of copyright law. . . . [T]he dispositive question is whether the copying goes to trivial or substantial elements.").

136. Meritorious cases tend to be settled in private through Writers Guild arbitration.

137. *See, e.g.*, *Stromback v. New Line Cinema*, 384 F.3d 283, 299 (6th Cir. 2004); *Flaherty, v. Filardi*, No. 03 Civ. 2167, 2007 U.S. Dist. LEXIS 69202, at *8–9 (S.D.N.Y. Sept. 19, 2007) (dismissing copyright claim to interim drafts of a published non-infringing final work as a matter of law); *Walker v. Time Life Films, Inc.*, 615 F. Supp. 430, 434–35 (S.D.N.Y. 1985) (denying request to discover drafts).

draft of the defendant's screenplay, in the hope that some earlier version of the work will disclose a greater resemblance to their own copyrighted work than the finished film does. Courts invariably deny these requests.¹³⁸ The reasons behind the denials provide an important insight into the structure of copyright law.

Courts refuse to entertain discovery with respect to early drafts of a non-infringing final work precisely because infringement requires at least some potential interference with the copyright owner's expectation of exclusivity. As noted in *Davis v. United Artists*, "the ultimate test of infringement must be the film as produced and broadcast, we do not consider the preliminary scripts."¹³⁹ Courts do not refuse to examine interim drafts merely because of judicial economy. As the Second Circuit noted in *Warner Bros., Inc. v. American Broadcasting Cos.*, "a defendant may legitimately avoid infringement by intentionally making sufficient changes in a work which would otherwise be regarded as substantially similar to that of the plaintiff's."¹⁴⁰ Likewise, in *See v. Durang*, the Ninth Circuit held "[t]he only discovery plaintiff suggests is the production of early drafts of defendant's play on the theory they might reflect copying from plaintiff's play that was disguised or deleted in later drafts. Copying deleted or so disguised as to be unrecognizable is not copying."¹⁴¹

The refusal of courts to entertain copyright infringement allegations in relation to unpublished drafts and preliminary scripts demonstrates the practical importance of a focus on expressive substitution. Because the copyright owner's rights are generally limited to the communication of their original expression to the public, even if it were not in the public domain, a filmmaker would be perfectly entitled to start with Jane Austen's *Emma* and

138. *See id.* at 435 (noting that courts routinely reject requests to consider earlier drafts of screenplays).

139. *Davis v. United Artists, Inc.*, 547 F. Supp. 722, 724 n.9 (S.D.N.Y. 1982) (citing *Fuld v. Nat'l Broad. Co.*, 390 F. Supp. 877, 882 n.4 (S.D.N.Y. 1975)); *see also Stromback*, 384 F.3d at 299 ("In deciding infringement claims, courts have held that only the version of the alleged infringing work presented to the public should be considered."); *Madrid v. Chronicle Books*, 209 F. Supp. 2d 1227, 1234 (D. Wyo. 2002) ("Since a court considers the works as they were presented to the public, discovery in this case . . . would be pointless.") (internal quotation marks omitted); *Walker*, 615 F. Supp. at 434 ("The Court considers the works as they were presented to the public.").

140. *Warner Bros. v. Am. Broad. Cos.*, 720 F.2d 231, 241 (2d Cir. 1983) (citing 3 NIMMER ON COPYRIGHT § 13.03[B] at 13-38.1 to 38.2; *Eden Toys, Inc. v. Marshall Field & Co.*, 675 F.2d 498, 501 (2d Cir. 1982); *Durham Indus. v. Tomy Corp.*, 630 F.2d 904, 913 & n.11 (2d Cir. 1980)). Courts addressing the question of intermediate copying in the software context have seen the matter slightly differently. *See infra*, note 172 and accompanying text.

141. *See v. Durang*, 711 F.2d 141 (9th Cir. 1983).

rework the plot over and over again until she comes out with *Clueless*.¹⁴² Intermediate scripts that never see the light of day do not communicate the author's original expression to the public and thus cannot constitute copyright infringement.

3. *The Implications of Computer Software and Other Functional Works Protected by Copyright Law*

Copyright protection for computer software has long been a source of controversy and disquiet.¹⁴³ Although the statutory definition of “literary works” in the Copyright Act is broad enough to include computer programs,¹⁴⁴ treating software as a work of literature presents something of a contradiction. The 1976 Copyright Act clearly states that copyright protection does not extend to any “process, system, [or] method of operation”¹⁴⁵ And yet, as made clear by a 1980 amendment to the Act, Congress intended that copyright protection would extend to computer programs.¹⁴⁶ The amendment defines a computer program as “a set of statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result.”¹⁴⁷ A “set of instructions” used “in order to bring about a certain result” appears to be the very essence of the “process, system, method of operation” exclusion under § 102(b).

With this contradiction in mind, it is hardly surprising that the general theory of copyright advanced in this Article—the centrality of expressive substitution—does not fit perfectly to software.¹⁴⁸ Users do not typically

142. CLUELESS (Paramount 1995). See Suzanne Ferriss, *Emma Becomes Clueless*, in JANE AUSTEN IN HOLLYWOOD 122 (Linda Troost & Sayre Greenfield eds. 2000).

143. See Jane C. Ginsburg, *Four Reasons and a Paradox: The Manifest Superiority of Copyright over Sui Generis Protection of Computer Software*, 94 COLUM. L. REV. 2559 (1994); Peter S. Menell, *An Analysis of the Scope of Copyright Protection for Application Programs*, 41 STAN. L. REV. 1045 (1989); Pamela Samuelson et al., *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308 (1994).

144. 17 U.S.C. § 101 (2010) (“literary works” includes works “expressed in words, numbers, or other verbal or numerical symbols or indicia”).

145. 17 U.S.C. § 102(b). Exclusive rights in processes and methods of operation are generally left to the patent system. See 35 U.S.C. § 101 (2010).

146. See *Apple Computer, Inc. v. Franklin Computer Corp.*, 714 F.2d 1240, 1247–49 (3d Cir. 1983) (reviewing legislative history); but see Pamela Samuelson, *CONTU Revisited: The Case Against Copyright Protection for Computer Programs in Machine-Readable Form*, 1984 DUKE L.J. 663.

147. Computer Software Copyright Act of 1980, Pub. L. No. 96-517, § 101, 94 Stat. 3028 (1980).

148. The same objections could be raised with respect to the copyright protection of architectural plans and the following discussion applies *mutatis mutandis* to that subject matter. The Berne Convention Implementation Act (1988) and the Architectural Works Copyright Protection Act (1990) recognize two separate forms of protection for architectural

copy copyrighted computer programs so that they can imbibe the artistry of the programmer's expression. Even if computer programs are to some extent expressive, they are predominantly functional.¹⁴⁹ The distinction between expressive and non-expressive uses is not intended to eviscerate copyright protection for computer software. As the preceding discussion makes clear, the rational justification for copyright is generally that it protects the author against expressive substitution. But the anomalous nature of computer software points to a different basis for attaching copyright protection and thus does not admit a defense of non-expressive use to the same extent. In sum, computer software (and other functional works that have been grafted onto copyright) should continue to be treated as exceptional—non-expressive use should not be regarded as a defense to ordinary acts of software piracy.¹⁵⁰

Combined with the idea-expression distinction, this brief review of the application of the tests for substantial similarity and fragmented non-literal similarity, and the refusal of courts to apply the author's reproduction right to intermediate drafts that never see the light of day, all point in the same direction: the copyright owner's exclusive rights are limited to the right to communicate the expressive aspects of her work to the public. This point is important because once it is understood that copyright's primary function is to protect the author from the threat of expressive substitution, the case in favor of non-expressive uses becomes almost self-evident. Standing alone, a non-expressive use carries no threat of expressive substitution and such uses should thus fall outside the scope of an author's entitlement.

D. ACTIVATING THE PRINCIPLE OF NON-EXPRESSIVE USE THROUGH FAIR USE

1. *Why Fair Use*

The preceding discussion concentrates on *why* we should recognize a general principle that non-expressive use is non-infringing; this Section turns

works, one for architectural plans and the other for structures based on such plans. For an overview, see 1 NIMMER ON COPYRIGHT § 2.08 (2012).

149. Pamela Samuelson, *supra* note 143, at 2315–18 (explaining that “the primary source of value in a program is its behavior, not its text”); Dennis S. Karjala, *Copyright Protection Of Computer Program Structure*, 64 BROOK. L. REV. 519, 532 (1998) (arguing that computer programs “are not like dictionaries or maps, which are useful only insofar as they supply information to human beings. A computer program is not intended to be ‘read’ or ‘understood’ by its target audience, let alone appeal to a user’s sense of esthetics.”).

150. However, as noted below, the non-expressive use analysis still provides a useful framework for understanding software reverse engineering cases. *See infra* notes 173–176 and accompanying text.

to the prescriptive implications of that principle, i.e., the question of *how* it should be recognized. The answer, in short, is that the reproduction of expressive copyrighted works for non-expressive uses requires context-specific review under the fair use doctrine for three reasons.

The first reason is simply that to hold otherwise would contradict the Copyright Act's plain language. Section 106(1) of the Act gives copyright owners the exclusive rights "to reproduce the copyrighted work in copies."¹⁵¹ Copies are defined as "material objects . . . in which a work is fixed . . . and from which the work *can* be perceived, reproduced or otherwise communicated."¹⁵² Thus, to make a *prima facie* infringing reproduction, one need only reproduce the work in a stable format such that it is *capable* of being perceived and used expressively. There is no express requirement in the Act that anyone actually perceives the work or uses it expressively.

The second reason, as already noted, is that blanket exclusion for non-expressive use would substantially undermine the legal protection of copyright's more irregular subject matter, such as computer software and architectural plans. Applying the principle of non-expressive use to anomalous copyright subject matter must be considered carefully. Rightly or wrongly, Congress has extended copyright protection to computer software and architectural plans to provide incentives for the development of these primarily functional objects.¹⁵³ Although computer programs are treated as expressive literary works, their expressive elements are secondary to the functional output of the program—i.e., what it actually does. In consequence, the everyday use of a computer program is non-expressive, but that does not suggest that copyright protection for software should be effectively dismantled. Instead, courts must exercise caution when dealing with anomalous copyright subject matter so as not to negate the very protection Congress intended.

The third reason not to adopt a *per se* rule with respect to non-expressive use is that in many contexts the concept is ambiguous. Like its subject matter equivalent, the idea-expression distinction, the line between expressive use and non-expressive use may often turn out to be a matter of context and degree. Where the validity of a defendant's claim that a particular

151. 17 U.S.C. § 106(1) (2010).

152. *Id.* § 101 (emphasis added).

153. *See* Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240, 1247 (3d Cir. 1983) (noting that "[a]lthough section 102(a) does not expressly list computer programs as works of authorship, the legislative history suggests that programs were considered copyrightable as literary works."); Architectural Works Copyright Protection Act, Pub. L. No. 101-650, 104 Stat. 5089, 5133 (1990).

use is non-expressive is contestable, courts may find that adopting a categorical rule that non-expressive uses are non-infringing simply shifts the argument's focus from substantive issues to questions of category definition.

For these three reasons, it is submitted that the principle of non-expressive use should be applied in the context of copyright's fair use doctrine and not as a freestanding defense to copyright infringement.

2. *Application to Fair Use*

This Section explores how the principle of non-expressive use should be (and, implicitly, is being) applied to the traditional four-factor fair use inquiry required under § 107 of the Copyright Act.¹⁵⁴

a) The "Purpose and Character" of Non-Expressive Uses

The non-expressive nature of the defendant's use is perhaps most relevant under the first fair use factor, "the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes."¹⁵⁵ Recognizing the copyright owner's exclusive rights as implicitly defined and limited in reference to expressive communication to the public makes sense of both expressive and non-expressive fair uses. Indeed, recognition of this overarching principle may be the key to rescuing the concept of transformative use from elastic imprecision.

According to the Supreme Court's most recent fair use decision, *Campbell v. Acuff-Rose*, the first factor turns primarily on:

[W]hether the new use merely supersedes the objects of the original creation . . . or instead adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message; it asks, in other words, whether and to what extent the new work is "transformative." . . . Although such transformative use is not absolutely necessary for a finding of fair

154. 17 U.S.C. § 107. The factors are:

(1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.

Id.

155. *Id.*

use . . . the goal of copyright, to promote science and the arts, is generally furthered by the creation of transformative works.¹⁵⁶

Traditionally, courts apply the concept of transformative use to new expressive uses that “provide social benefit, by shedding light on an earlier work, and, in the process, creat[e] a new one.”¹⁵⁷ Transformative use is most obvious when the work is itself transformed; however, in many cases courts have held that the mere recontextualization of a copyrighted work from one expressive context to another is sufficient to sustain a finding of fair use—the work itself need not be altered.¹⁵⁸

Understanding the rationale for transformative use is the key to grasping the link between transformative use and non-expressive use. The privileged status of transformative uses under the fair use doctrine allows for the creation of new works from old. This is not a sufficient explanation, however, because other doctrinal levers, such as a narrower understanding of the author’s exclusive right to make derivative works, could achieve the same effect.¹⁵⁹ Beyond a simple enthusiasm for new works based on the copyrighted work, courts accord special status to transformative uses because they do not substitute for the author’s original expression—they do not merely supersede the objects of the original creation.¹⁶⁰ Because of this special status, the greater the extent of transformation, the less significant other factors weighing against fair use will become.¹⁶¹

Cognizant of the Supreme Court’s focus on transformative uses, some courts have simply equated non-expressive with transformative. In *Perfect 10, Inc. v. Amazon, Inc.*, the court held that Google’s use of thumbnails in its Internet search engine “may be more transformative than a parody because a search engine provides an entirely new use for the original work, while a

156. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994) (citations and internal quotation marks omitted); see also Pierre N. Leval, Commentary, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1111 (1990) (“I believe the answer to the question of justification turns primarily on whether, and to what extent, the challenged use is transformative.”).

157. *Campbell*, 510 U.S. at 579.

158. See, e.g., *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 609–10 (2d Cir. 2006) (holding that use of promotional posters in a rock biography was “a purpose separate and distinct from the original artistic and promotional purpose for which the images were created”); *Mattel, Inc. v. Walking Mountain Prods.*, 353 F.3d 792, 796–98, 800–06 (9th Cir. 2003) (concluding that photos parodying Barbie by depicting “nude Barbie dolls juxtaposed with vintage kitchen appliances” was a fair use).

159. See, e.g., 17 U.S.C. § 106(2).

160. See, e.g., *Campbell*, 510 U.S. at 579.

161. *Id.*

parody typically has the same entertainment purpose as the original work.”¹⁶² This scenario seems to be stretching the concept of transformation beyond its natural utility. It would be better to recognize uses that do not relate to the expressive appeal of a work may find favor under the first fair use factor—whether they qualify as transformative in the expressive sense or not.

By construction, the more non-expressive the use of a copyrighted work is, the less it substitutes for the author’s original expression.¹⁶³ As such, courts should regard primarily non-expressive uses as equivalent (but not identical) to highly transformative uses—their “purpose and character” is such that they do not merely supersede the objects of the original creation.¹⁶⁴ In addition, the same logic that dictates that the more transformative a work is, the less significant the other factors become, also applies to non-expressive uses.¹⁶⁵

b) Non-Expressive Use and Commercial Fair Use

While considering the “purpose and character of the use” under the first factor, courts are instructed to consider “whether such use is of a commercial nature or is for nonprofit educational purposes.”¹⁶⁶ The status of commercial fair use has proved to be confusing in the fair use case law, in part because it is so closely linked with the question of market substitution under the fourth factor.¹⁶⁷ Even if commercial entities develop and maintain copy-reliant technologies such as search engines, this does not weaken their claim to fair use.¹⁶⁸ If a use is non-expressive, its commercial or noncommercial nature is

162. *Perfect 10, Inc. v. Amazon, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007) (holding further that “even making an exact copy of a work may be transformative so long as the copy serves a different function than the original work” (citing *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818–19 (9th Cir. 2003)).

163. The analysis in this Section is subject to the caveat regarding computer software and other quasi-functional works discussed in Section III.C.3, *supra*.

164. *See Campbell*, 510 U.S. at 583.

165. *See id.* at 579.

166. 17 U.S.C. § 107(1) (2010).

167. *Sag, Predicting Fair Use*, *supra* note 22, at 58–61. Indeed, the Ninth Circuit’s approach to commerciality in *Napster* defines the concept exclusively in terms of market substitution. *See A&M Records v. Napster*, 239 F.3d 1004, 1015 (9th Cir. 2001) (holding that “commercial use is demonstrated by a showing that repeated and exploitative unauthorized copies of copyrighted works were made to save the expense of purchasing authorized copies”).

168. This assessment is reinforced by recent empirical analysis of fair use cases in U.S. district courts, *Sag, Predicting Fair Use*, *supra* note 22, at 77 (finding that there is no evidence that commercial use plays any objectively ascertainable role in determining the outcome of fair use cases). Non-commercial entities such as universities may have an especially strong claim to fair use for reasons *related to* their non-commercial status, but *not because of the status itself*. For example, copying by a university for the purposes of research or education may be

irrelevant because non-expressive uses do not substitute for the author's original expression.¹⁶⁹

c) Non-Expressive Use and "Amount and Substantiality"

The degree that a use is non-expressive is also significant in terms of the third fair use factor, "the amount and substantiality of the portion used in relation to the copyrighted work as a whole."¹⁷⁰ Far from being linear or arithmetic in nature, proper application of the third factor is contingent upon the purpose and the effect of the defendant's use.¹⁷¹ Instead of relying on a mechanical quantification of the *amount* of the original work used, the third factor asks courts to assess how much of the *value* of the original work is present in the allegedly infringing work.¹⁷² Accordingly, the extent to which a use is non-expressive plays a vital role in the assessment of the third fair use factor. A non-expressive use does not generally substitute for the expressive value of the author's original expression, and therefore courts should view it as qualitatively insignificant under the third factor, even if it involves literal copying of an entire work.

This insight helps us make sense of the superficial conflict between Hollywood cases alleging intermediate copying and analogous Silicon Valley cases.¹⁷³ In cases involving motion pictures, courts have refused to apply the author's reproduction right to allegedly infringing intermediate drafts of screenplays. However, courts addressing the question of intermediate copying in the software context have seen the matter slightly differently.¹⁷⁴ In software reverse engineering cases, courts appear to take the allegation of infringement via intermediate copying seriously as a potential basis for infringement.¹⁷⁵ This difference is best explained by the anomalous nature of

less likely to have a market effect or may generate positive externalities, which make efficient licensing less likely.

169. See *supra* note 140 and accompanying text (noting the caveat relating to anomalous copyright subject matter such as computer software).

170. 17 U.S.C. § 107(3). This inquiry can be traced back to Justice Story's original formulation of the fair use doctrine in *Folsom v. Marsh*, 9 F. Cas. 342 (C.C. Mass. 1841) (No. 4,901). In that case, Justice Story was concerned to protect the "chief value of the original work" against the extraction of its "essential parts" through the mere "facile use of scissors" or its intellectual equivalent. *Id.* at 345.

171. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 586–87 (1994) ("the extent of permissible copying varies with the purpose and character of the use").

172. See Matthew Sag, *God in the Machine, A New Structural Analysis of Copyright's Fair Use Doctrine*, 11 MICH. TELECOMM. & TECH. L. REV. 381, 391 (2005).

173. The terms Hollywood and Silicon Valley are used representationally.

174. See, e.g., *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1518–19 (9th Cir. 1992).

175. *Id.* at 1519. The *Sega* court found that:

computer software itself. Although software is protected under the expressive regime of copyright law, the value of software that the law is protecting relates to the function or behavior of the code, not to its expression. In contrast, a screenplay has no behavioral value beyond the communication of the author's expression to the public. Thus it makes sense that in film cases allegations of intermediate copying would be categorically dismissed, whereas in software cases the courts would take a more contextual approach and review the allegation as a question of fair use.¹⁷⁶

Returning to the third factor itself, the reverse engineering cases nicely illustrate the contention that non-expressive uses do not generally substitute for the value of the work. In *Sony v. Connectix*, for example, the court acknowledged that Connectix had copied an entire section of Sony's software multiple times; however, it concluded "in a case of intermediate infringement when the final product does not itself contain infringing material, this factor is of very little weight."¹⁷⁷

d) The Market Effect of Non-Expressive Uses

The fourth fair use factor is "the effect of the use upon the potential market for or value of the copyrighted work."¹⁷⁸ Of course, the question of market effect risks collapsing into tautology because every use by a defendant represents something that could, in theory, be licensed to the defendant if the court rules that such use is not fair use. But courts avoid this circular reasoning by limiting the abstract market to a market that is cognizable under copyright. The market harms that courts refuse to recognize illustrate again that the copyright owner's exclusive rights are limited to the communication of their original expression to the public. The case law indicates that courts exclude consideration of market effects that do not arise from expressive substitution.

[I]ntermediate copying . . . may infringe the exclusive rights granted to the copyright owner in section 106 of the Copyright Act regardless of whether the end product of the copying also infringes those rights. If intermediate copying is permissible under the Act, authority for such copying must be found in one of the statutory provisions to which the rights granted in section 106 are subject.

Id.

176. However, the reverse engineering cases all find that the practice is fair use, suggesting that future courts might invoke a per se analysis for the sake of judicial economy. *See infra* note 179.

177. *Sony Computer Entm't, Inc. v. Connectix Corp.*, 203 F.3d 596, 606 (9th Cir. 2000) (internal quotation marks omitted).

178. 17 U.S.C. § 107(4) (2010).

In *Campbell*, the Supreme Court quite plainly differentiated the copyright owner's general economic interests from the limited protection afforded by copyright:

[W]hen a lethal parody, like a scathing theater review, kills demand for the original, it does not produce a harm cognizable under the Copyright Act. Because parody may quite legitimately aim at garroting the original, destroying it commercially as well as artistically, the role of the courts is to distinguish between biting criticism that merely suppresses demand and copyright infringement, which usurps it.¹⁷⁹

Just as *Campbell* recognizes that criticism is outside of the copyright owner's protectable sphere of interest, the reverse engineering cases recognize that the copyright owner has no protectable interest in preventing the copying of unprotectable expression and ideas buried within its object code. Courts have consistently held that making unauthorized copies of a computer program, as a necessary step in reverse engineering, is fair use.¹⁸⁰ For example, in *Sony v. Connectix*, the Ninth Circuit held that although the defendant's Virtual Game Station console directly competed with Sony in the market for platforms capable of playing Sony Playstation games, the Virtual Game Station was a "legitimate competitor" in that market.¹⁸¹ The court concluded that Sony's desire to control the market for gaming platforms was understandable but that "copyright law . . . does not confer such a monopoly."¹⁸²

The treatment of parody and reverse engineering illustrates the exclusion of market effects that do not arise from expressive substitution. This rationale is implicit in *Campbell* where the Court notes "[p]eople ask for

179. *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 591–92 (1994) (internal quotation marks and citations omitted).

180. *See, e.g., Sony Computer Entm't*, 203 F.3d at 606, *cert. denied*, 531 U.S. 871 (2000) (holding that Connectix's copying of Sony's copyrighted basic input-output system (BIOS) during reverse engineering, used by Connectix to develop a software program that emulates the functioning of the Sony PlayStation console for regular computers, was fair use); *Atari Games Corp. v. Nintendo of Am., Inc.*, 975 F.2d 832, 842–43 (Fed. Cir. 1992) (observing that Atari's reverse engineering of Nintendo's 10NES program would have been a fair use of the program, except that Atari did not possess an authorized copy of the work); *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510, 1520 (9th Cir. 1992) (holding that Accolade's reverse engineering of Sega's video game programs in order to figure out how to make its own games compatible with Sega's Genesis system is a fair use); *see also* David A. Rice, *Copyright and Contract: Preemption After Bowers v. Baystate*, 9 ROGER WILLIAMS U. L. REV. 595, 601 n.19 (2004) (collecting cases). Circumventing encryption for the purpose of reverse engineering is also allowed under the safe harbor provisions of the DMCA. *See* 17 U.S.C. § 1201(f).

181. *Sony Computer Entm't*, 203 F.3d at 607; *see also* *Sega*, 977 F.2d at 1522–23.

182. *Sony Computer Entm't*, 203 F.3d at 607 (emphasis added); *see also* *Sega*, 977 F.2d at 1523–24.

criticism, but they only want praise.”¹⁸³ Thus, “the unlikelihood that creators of imaginative works will license critical reviews or lampoons of their own productions removes such uses from the very notion of a potential licensing market.”¹⁸⁴ This rationale is explicit in the reverse engineering cases. From the beginning of its decision in *Sony v. Connectix*, the court emphasized the importance of the idea-expression distinction: “[w]e are called upon once again to apply the principles of copyright law to computers and their software, to determine what must be protected as expression and what must be made accessible to the public as function.”¹⁸⁵ Consistent with its decision in *Sega Enterprises v. Accolade, Inc.*,¹⁸⁶ the Ninth Circuit held in *Sony v. Connectix* that intermediate copying of software is fair use if the copying was necessary to gain access to the software’s functional elements.¹⁸⁷ The court based its ruling firmly on the importance of maintaining the idea-expression distinction: “[w]e drew this distinction because the Copyright Act protects expression only, not ideas or the functional aspects of a software program Thus, the fair use doctrine preserves public access to the ideas and functional elements embedded in copyrighted computer software programs.”¹⁸⁸

In the case of expressive uses such as parody, and non-expressive uses such as reverse engineering, courts have consistently held that the protection that copyright affords is limited to certain cognizable markets. Transformative expressive uses do not usually affect the market in any relevant sense because the second author’s expression does not substitute for that of the original author. The absence of any cognizable market effect is even more apparent in cases of non-expressive use because, to the degree that a particular use is non-expressive, it has no potential substitution effect on a cognizable copyright market.

As established earlier in this Part, the copyright owner’s exclusive rights typically hinge upon the communication of original expression to the public. Acts of copying that do not communicate the author’s original expression to the public should not generally be held to constitute copyright infringement. The most appropriate method of doctrinal incorporation of the principle of non-expressive use is through the fair use doctrine. The role of expressive

183. *Campbell*, 510 U.S. at 592 (quoting SUMMERSET MAUGHAM, OF HUMAN BONDAGE 241 (Penguin ed. 1992)).

184. *Id.*

185. *Sony Computer Entm’t*, 203 F.3d at 598.

186. *Sega*, 977 F.2d at 1510.

187. *Sony Computer Entm’t*, 203 F.3d at 607.

188. *Id.* at 603 (citing *Sega*, 977 F.2d at 1510).

substitution is not merely compatible with the fair use doctrine; more accurately, expressive substitution is necessary to make sense of much existing case law. It is unrealistic to attempt to reduce the entirety of fair use jurisprudence into any one coherent principle. Nonetheless, the general proposition that the doctrine favors acts of copying unlikely to substitute for the copyright owner's original expression explains the majority of cases. Like transformative expressive uses, primarily non-expressive uses should generally be classified as fair uses because, by their very nature, they do not substitute for the author's original expression. Accordingly, like transformative use, non-expressive use should be favored under the first, third, and fourth factors—such uses are non-substitutive in “purpose and character,” appropriate a qualitatively insignificant proportion of the value of the copyright owner's original expression, and produce no cognizable market effect under the fourth factor.¹⁸⁹

IV. CONCLUSION: UNLEASH THE MACHINES

Digital technology offers powerful tools for organizing, analyzing, and searching through an otherwise overwhelming sea of information. The legality of these tools has generally been accepted in the purely online context of text-based and visual search engines and the context of software enabled plagiarism detection systems.¹⁹⁰ The library digitization debate brings the same issue to a new context: printed books.

The Authors Guild's campaign against the Google book search initiative came to an abrupt halt with the proposal of a class action settlement in 2008, followed by an Amended Settlement Agreement in 2009. That agreement has since been rejected by the supervising court and the legality of Google's initiative is still disputed by many authors and publishers. Google has provided electronic versions of millions of library books to the university libraries that made the paper copies initially available. Those universities must now determine how, if at all, they should use this resource. In 2008, several

189. As is so often the case, the second statutory factor does not appear to have much bite in the context of non-expressive uses, and thus does little to “separat[e] the fair use sheep from the infringing goats.” *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 586 (1994). See Sag, *Predicting Fair Use*, *supra* note 22.

190. *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630 (4th Cir. 2009); *Perfect 10, Inc. v. Amazon, Inc.*, 508 F.3d 1146 (9th Cir. 2007); *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003); *Field v. Google, Inc.*, 412 F. Supp. 2d 1106, 1117–19 (D. Nev. 2006).

universities agreed to combine their digital collections in a shared repository called the HathiTrust.¹⁹¹

In September 2011, the Authors Guild announced that it was suing five universities and the HathiTrust for the “systematic, concerted, widespread and unauthorized reproduction and distribution of millions of copyrighted books”¹⁹² The Guild objects to the universities’ plan to distribute works for which they have been unable to locate the copyright owner, i.e., orphan works.¹⁹³ Implausibly, the Guild stakes the claim that libraries are not entitled to fair use under § 107 of the Copyright Act because libraries are the beneficiary of a more limited exemption under § 108.¹⁹⁴ Whether the limited reproduction and distribution of orphan works is permitted by fair use is a crucial question, but it is not the subject of this Article. Instead, this Article addresses the other aspect of the Guild’s claim, the assertion that even library digitization restricted to enabling data analysis constitutes “one of the largest copyright infringements in history”¹⁹⁵

The HathiTrust aims to develop and facilitate the development of data mining and analysis of its digital collection.¹⁹⁶ This activity would have qualified as “non-consumptive research” under the now defunct Amended Settlement Agreement (“ASA”).¹⁹⁷ “Non-consumptive research” as defined in the ASA is a form of non-expressive use as the term is used in this Article. According to the Authors Guild, in the absence of a class action settlement

191. The HathiTrust includes material provided by Google, the Internet Archive, Microsoft, and the universities themselves. See *Our Digital Library*, HATHITRUST DIGITAL LIBRARY, http://www.hathitrust.org/digital_library (July 17, 2012).

192. Authors Guild, *Australian Society of Authors, Quebec Writers Union Sue Five U.S. Universities*, AUTHORS GUILD BLOG (Sept. 12, 2011), <http://blog.authorsguild.org/2011/09/12/authors-guild-australian-society-of-authors-quebec-writers-union-sue-five-u-s-universities-2/>. In October 2012, the district court released its opinion in *Authors Guild, Inc. v. HathiTrust*, 11 CV 6351 HB, 2012 WL 4808939 (S.D.N.Y. Oct. 10, 2012).

193. First Amended Complaint, *Authors Guild v. HathiTrust*, No. 11 Civ. 6351, (S.D.N.Y. Sept. 12, 2011).

194. *Id.*

195. *Id.* ¶ 7. Paragraph 68 of the Amended Complaint also states:

[U]sers may search and identify bibliographic information (title, author, subject, ISBN, publisher, and year of publication) for the works contained in the HDL. HathiTrust also permits all users to search the entire text of all works in the HDL (including public domain and in-copyright works) to determine the number of times and page location(s) of any keyword or phrase found in a book.

Id. ¶ 68.

196. *Functional Objectives*, HATHITRUST DIGITAL LIBRARY (Nov. 5, 2010), <http://www.hathitrust.org/objectives>.

197. ASA, *supra* note 23, § 1.93.

or some express authorization by copyright owners, the creation of systems for the automated analysis of library books constitutes copyright infringement.¹⁹⁸ If this is correct, then the non-expressive use of copyrighted works will be impeded: the large number of permissions required and the difficulty of locating and identifying the relevant interests makes right-clearance on the scale of millions of works implausible.¹⁹⁹

Where large-scale electronic text collections are available, advances in computational power and a proliferation of new text mining and visualization tools offer scholars of the humanities the chance to do what biologists, physicists, and economists have been doing for decades—analyze data.²⁰⁰

Scholars in the “Digital Humanities” believe that text mining and the computational analysis of text are vital to the progress of human knowledge in the “Information Age.” The potential of these non-expressive uses of text has already been made apparent in the life sciences where researchers use a variety of text-mining tools to accelerate the identification of relevant research across disparate fields and to suggest hitherto unseen correlations or associations such as protein-protein interactions and gene-disease associations.²⁰¹

Similar breakthroughs are on the horizon in the humanities. Traditionally, literary scholars have relied upon the close and often anecdotal study of select works. Modern computing power and the mass-digitization of texts now permits investigation of the larger literary record.

Literary analyses of digitized collections are at the core of Digital Humanities research. Large scale quantitative projects such as those being undertaken at the Stanford Literary Lab are unearthing previously unknowable information about individual works, genres, and even entire eras.²⁰² Digitization enhances our ability to process, mine, and ultimately

198. First Amended Complaint, *supra* note 193.

199. Imagine someone other than a phone company trying to write a new telephone book and having to ask every household for permission.

200. This paragraph and remainder of the text were written in parallel with two amicus briefs, one in *Authors Guild v. HathiTrust* and one in *Authors Guild v. Google*. Matthew Jockers, Jason Schultz, and I jointly authored these briefs. We were assisted by many people, most notably, David Hansen and Ana Enriquez.

201. Sophia Ananiadou, Douglas B. Kell & Jun-ichi Tsujii, *Text mining and its potential applications in systems biology*, 24:12 TRENDS IN BIOTECHNOLOGY 571 (2006) (citing Christian Blaschke et al., *Information extraction in molecular biology*, 3 BRIEF. BIOINFORM. 154 (2002); Toshihide Ono et al., *Automated extraction of information on protein-protein interactions from the biological literature*, 12 BIOINFORMATICS 155 (2001)).

202. The Stanford Literary Lab discusses, designs, and pursues literary research of a digital and quantitative nature. *About*, STANFORD LITERARY LAB, <http://litlab.stanford.edu/> (last visited June 18, 2012).

better understand individual texts, the connections between texts, and the overall evolution of literary language. As Matthew Jockers explains, by exploring the literary record writ large, researchers can better understand the context in which individual texts exist and thereby better understand those individual texts.²⁰³ As Franco Moretti has further noted, “a field this large cannot be understood by stitching together separate bits of knowledge about individual cases, because it isn’t a sum of individual cases: it’s a collective system, that should be grasped as a whole”²⁰⁴ Grasping a system as a whole is not possible without the ability to make non-expressive uses of digitized text. For some, the possibility of mining huge digital archives has been a major catalyst for changing the very conception of humanities research. For others, it is a useful tool for testing old theories or suggesting new areas of inquiry.

Researchers in Information Retrieval frequently use text-mining and computer-aided classification to identify and retrieve relevant documents. Using similar techniques, researchers in the Digital Humanities use text mining and computer-aided classification to identify and retrieve relevant texts, often found in unlikely places. This enables researchers in the humanities to expand their traditional study of a few, canonical works to a study of any one of the several million books in the larger archive of literary history, an archive that has hitherto remained hidden because of the limitations of human reading. Thus, non-expressive use leads to additional expressive use and thus expands the audience (and the potential market) for individual works.²⁰⁵

Moreover, digitization also allows scholars to reimagine the relationships between texts. For example, the Google Ancient Places project links the text of public domain books such as *Gibbon’s Decline and Fall of the Roman Empire* to a map of the ancient world.²⁰⁶ The interface allows the user to browse the books, including the full text, at the same time as she browses a map. The places are marked on the map and hyperlinked. Again, the map itself is a non-expressive use of the underlying texts, but such use may well still lead to

203. JOCKERS, *supra* note 116.

204. MORETTI, *supra* note 115, at 4.

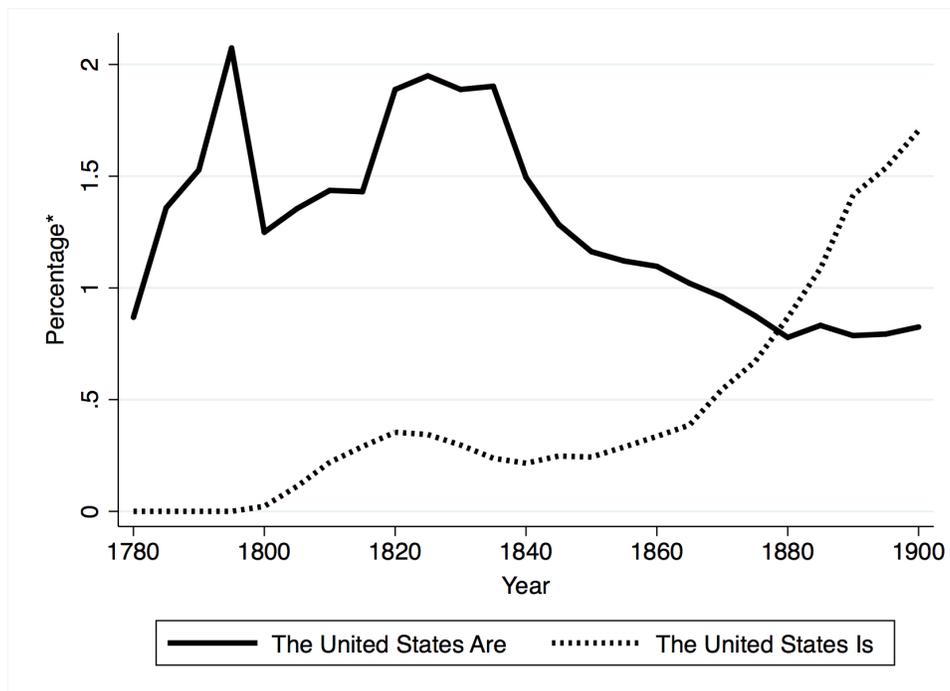
205. For example, Matthew Jockers used text-mining and computer-aided classification to identify an overlooked tradition of whaling fiction predating (and arguably informing) Melville’s writing of *Moby Dick*. See JOCKERS, *supra* note 116.

206. *About*, GOOGLE ANCIENT PLACES, <http://googleancientplaces.wordpress.com/about/> (last visited July 17, 2012).

additional expressive use and expansion of the audience—and, again the potential market—for individual works.²⁰⁷

The Google Ngram tool provides a simple example of such non-expressive use. The comparison of the frequency with which authors refer to the United States as a single entity (“is”) versus a collection of individual states (“are”) is only possible with a digitized archive of significant size and coverage.²⁰⁸

Figure 5: Google Ngram Visualization Comparing Frequency of “The United States is” to “The United States are”



207. In a similar vein, researchers at Stanford University have mapped thousands of letters exchanged during the Enlightenment and have pieced together how these individual networks fit into a complete whole they refer to as the “Republic of Letters.” See *Mapping the Republic of Letters*, <https://republicofletters.stanford.edu/> (last visited June 20, 2012). One such visualization yields the surprising insight that although Voltaire admired England for its tolerance, freedom and political institutions, surprisingly few letters actually went to England. See Patricia Cohen, *Digital Keys for Unlocking the Humanities’ Riches*, N.Y. TIMES, Nov. 17, 2010, at C1.

208. *Google Books Ngram Viewer*, GOOGLE.COM, <http://books.google.com/ngrams> (last visited June 30, 2012). Figure 5 is a reconstruction of data generated using Google Ngram, sampled at 5-year intervals. The y-axis is scaled to 1/100,000 of a percent, such that 1=0.00001%. This particular ngram can be reproduced as follows: http://books.google.com/ngrams/graph?content=The+United+States+is%2C+The+United+States+are&year_start=1780&year_end=1900&corpus=5&smoothing=10.

Note that metadata produced in this visualization was only possible because the entire contents of the relevant books had been digitized. But note also that not a single sentence of the underlying books has been reproduced in the finished product. This kind of non-expressive use may add to our understanding, appreciation, and enjoyment of copyrighted works, but since it does not allow for the underlying works to be reconstructed, it could hardly be said to substitute for the originals.²⁰⁹

Google Ngram is just the tip of an emerging iceberg.²¹⁰ In a forthcoming book *Macroanalysis: Digital Methods and Literary History*,²¹¹ Professor Jockers draws on a corpus of nineteenth century novels to demonstrate how literary style changes over time. By studying word frequencies, syntactic patterns, and thematic markers in the context of metadata about author nationality, author gender, and historical time period, this kind of work opens up literary study to an entirely new perspective. Thus, in the larger context of the digital archive, Jockers is able to identify both the trendsetters and the outliers. Text mining and computational analysis can lead to surprising results. For example, Jockers demonstrates that Harriet Beecher Stowe's fiction is far more similar to the work of male authors of her generation than to the typically female-authored works of sentimental fiction among which her work is generally categorized.

The macro analysis of text archives has the potential to yield specific insights into literary historical questions, such as the historic place of individual texts, authors, and genres in relation to a larger literary context; literary patterns and lexicons employed over time, across periods, within regions, or within demographic groups; the cultural and societal forces that impact literary style and the evolution of style; the waxing and waning of

209. For additional examples of the use of Ngram, see, for example, Jean-Baptiste Michel, et al., *Quantitative Analysis of Culture Using Millions of Digitized Books*, 331 SCIENCE 176 (2011) (a study of linguistic and cultural changes in over five million digitized books) available at <http://www.sciencemag.org/content/331/6014/176>.

210. The toolkit available to Digital Humanities researchers is becoming increasingly sophisticated. See, e.g., TAPOR <http://www.tapor.ca/> (last visited June 30, 2012) (tools to map word usage over time, including peaks, density, collocations, and types); Andrew Kachites McCallum, *MALLET: A Machine Learning for Language Toolkit*, <http://mallet.cs.umass.edu/> (last visited June 30, 2012) (a Java-based package for statistical natural language processing, document classification, clustering, topic modeling, information extraction, and other machine learning applications to text); *MONK: Metadata Offer New Knowledge*, <http://www.monkproject.org/> (last visited June 30, 2012) (a digital environment designed to help humanities scholars discover and analyze patterns in the texts); SEASR: THE SOFTWARE ENVIRONMENT FOR THE ADVANCEMENT OF SCHOLARLY RESEARCH, <http://www.seasr.org> (last visited June 30, 2012).

211. UIUC Press (forthcoming 2013).

literary themes; and the tastes and preferences of the literary establishment and whether those preferences correspond to general tastes and preferences.²¹² *Realizing that potential requires analytical tools and capabilities and access to digitized texts.*

And yet, today's digital-minded literary scholar is shackled in time. In the absence of a policy allowing non-expressive use of copyrighted material, literary scholars, historians, and other humanists are all destined to become nineteenth-centuryists: slaves not to history, but to the public domain. To do their work thoroughly and completely—to study literary history, cultural history, and the human record writ large—these scholars simply must have access to the source material of literary, cultural, and human history. This history does not and should not end in 1923.²¹³

One of the aims of this Article is to disentangle the library digitization issue for the purposes of data analysis from the broader orphan works debate. There is no orphan works problem for library digitization-search because the copyright owners are not implicated by digitization for the purpose of non-expressive use. The distinction between expressive and non-expressive *works* is already well recognized in copyright law as the gatekeeper to copyright protection. As this Article has shown, the same distinction should generally be made in relation to potential *acts* of infringement. Preserving the functional force of the idea-expression distinction in the digital context requires courts to conclude that copying for purely non-expressive purposes, such as the automated extraction of data, are not infringing. Like transformative uses, such as parody and criticism, non-expressive uses should generally be classified as fair use because, by their very nature, they do not substitute for the author's original expression.

The legal status of actual copying for non-expressive uses was not a burning issue before digital technology. Outside the context of reading machines like search engines, plagiarism software, and the like, courts have quite reasonably presumed that every copy of an expressive work is for an expressive or consumptive purpose. The issue is now, however, squarely before the courts and should be addressed. To apply the words of the Ninth Circuit Court of Appeals in *Sony v. Connectix* in a different context, “[courts] are called upon once again to apply the principles of copyright law to [the use of] computers . . . , to determine what must be protected as expression and what must be made accessible to the public”²¹⁴

212. See JOCKERS, *supra* note 116.

213. *Id.*

214. *Sony Computer Entm't, Inc. v. Connectix Corp.*, 203 F.3d 596, 598 (9th Cir. 2000).

The idea-expression distinction protects the author's legitimate interest in her work while guaranteeing others the breathing space to supplement, reuse, or reinterpret the facts and ideas embodied in the work. A similar distinction should be applied to enable the non-expressive use of copyrighted works in the age of reading machines, even if those machines reproduce the text as a step in the analytical process.

