RISING ABOVE LIABILITY: THE DIGITAL SERVICES ACT AS A BLUEPRINT FOR THE SECOND GENERATION OF GLOBAL INTERNET RULES

Martin Husovec[†]

ABSTRACT

Twenty-five years ago, in 1998, the United States Congress developed a blueprint for the global regulation of the internet. Section 512 of the Digital Millennium Copyright Act (DMCA) recognized that user-generated content will be crucial to most digital services and offered up-front assurances from liability to some providers subject to conditions. What started as a sectorial conditional immunity system in copyright law was immediately scaled up into an all-encompassing horizontal rulebook in the European Union through the E-Commerce Directive (ECD) in 2000—recently updated into the Digital Services Act (DSA). The last two decades have largely validated the DMCA's conditional immunity as a feasible baseline approach to the regulation of internet communications that power global exchanges of ideas, goods, and services. However, the conditional immunity model has its limits. It was not designed to offer a complex solution for new challenges. The DSA is the first comprehensive attempt to create a second generation of rules for digital services that rely on user-generated content. Unlike previous sectorial initiatives, its approach is sweepingly horizontal. The DSA requires some level participation from both state and non-state institutions for its system of checks and balances to work, and some of its solutions can be "too European." However, the principles behind the DSA could be useful in other jurisdictions—perhaps even in the United States. The United Kingdom, which is currently developing its own set of post-Brexit rules, continues to build on some of the same principles as the DSA.

TABLE OF CONTENTS

DOI: https://doi.org/10.15779/Z38M902431

† Associate Professor of Law at London School of Economics and Political Science (LSE). I am grateful for feedback from editors and reviewers which enriched this Article. The mistakes are solely mine.

I. INTRODUCTION

Twenty-five years ago, in 1998, the United States Congress developed a blueprint for the global regulation of the internet. Section 512 of the Digital Millennium Copyright Act[1] (DMCA) recognized that user-generated content will be crucial to most digital services and offered up-front assurances from liability to some providers subject to conditions. What started as a sectorial, conditional immunity system in copyright law was immediately scaled up into an all-encompassing horizontal rulebook in the European Union through the E-Commerce Directive (ECD) in 2000[2]—recently updated into the Digital Services Act (DSA).[3] The two jurisdictions inspired many other countries to start granting conditional immunity—liability exemptions that require at least providers' knowledge of others' actions to expose them to liability for those actions.[4]

1. 17 U.S.C. § 512.

2. Directive 2000/31 of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, O.J. (L 178), 1–16 (commonly and hereinafter referred to as the E-Commerce Directive).

3. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act), O.J. (L 277) 1 EU.

4. *See, e.g.*, The Information Technology Act, 2000, § 79 (Indian law covering conduit and hosting services); Information Technology Framework Act, R.R.Q. 2001, c C-1.1 (Canadian law covering hosting and search engine services); Lei No. 12.965, de 23 de Abril de 2014, Diario Oficial da Uniao [D.O.U] de 24.04.2014 (Brazilian law covering conduit and hosting services). Attempts to introduce exemptions sometimes took different turns; for example, South Korean liability exemptions were turned into liability norms. Act on Promotion of Information and Communications Network Utilization and Information Protection, art. 44-2, *translated in* Korea Legislation Research Institute's online database, https://elaw.klri.re.kr/eng_service/main.do (search required).

Unlike its older sister, § 230[5] of the Communication Decency Act (CDA) adopted in 1996,[6] § 512 of the Digital Millennium Copyright Act is not widely credited as having created the internet.[7] Yet, upon closer look, while § 230 of the CDA might continue to guarantee the internet as we know it in the legal system of the United States, it is the DMCA's model that continues to run the internet globally. For many countries for which § 230 offers a constitutionally unacceptable immunity model for application-layer services,[8] the DMCA offers a more acceptable version. The DMCA-style conditional immunity is therefore also increasingly present in bilateral trade agreements.[9] If we ever witness international harmonization on the issue, this type of conditional immunity model is probably more likely to prevail.[10]

In Europe, conditional immunity was powerfully used in the infancy of new digital markets to unite countries under one set of rules. The ingenuity of

5. 47 U.S.C. § 230.

6. Today's broad reading of § 230 CDA is a result of the judicial reading in Zeran v. Am. Online Inc., 129 F.3d 327 (4th Cir. 1997) that rejected a narrower understanding that would allow distributors to be held liable based on their knowledge of illegal content, and Batzel v. Smith, 333 F.3d 1018, 1033 (9th Cir. 2003) that allowed providers to participate in the selection process to a limited degree.

7. Kosseff makes this point most forcefully in his book. *See generally* JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET (Cornell Univ. Press 2019).

8. In the European legal system, denial of remedy in cases like *Batzel*, 129 F.3d 327 or *Zeran*, 333 F.3d 1018 would constitute violation of Article 8 of the European Convention on Human Rights (ECHR), which is evident in cases like K.U. v. Finland, App. No. 2872/02 (Dec. 2, 2008), https://hudoc.echr.coe.int/fre?i=001-89964; Delfi AS v. Estonia, App. No. 64669/09, ¶ 110 (Jun. 16, 2015), https://hudoc.echr.coe.int/app/conversion/pdf/?library=ECHR&id=001-155105&filename=001-155105.pdf; and most recently Sanchez v. France, App. No. 45581/15, ¶ 162 (Sept. 2, 2021), https://hudoc.echr.coe.int/fre?i=001-211599 ("While the Court acknowledges that important benefits can be derived from the internet in the exercise of freedom of expression, it has also found that the possibility of imposing liability for defamatory or other types of unlawful speech must, in principle, be retained, constituting an effective remedy for violations of personality rights").

9. Daniel Seng, *The State of the Discordant Union: An Empirical Analysis of DMCA Takedown Notices*, 18 VA. J. L. & TECH. 369, 374 (2014) ("[T]he DMCA safe harbors have indeed gone global. And the world has embraced the DMCA."). Seng lists some of the FTAs at pages 373–75.

10. *See* WTO, WTO Electronic Commerce Negotiations Updated Consolidated Negotiating Text, WTO INF/ECON/62/Rev.2 (Sept. 2021) (limiting liability through Article B.1(2)). However, even Article 19.17.2 of the Canada-US-Mexico Trade Agreement, which contains a provision inspired by Section 230 of the CDA, was interpreted by Canadian courts as permitting a Canadian DMCA-inspired notice-based liability exemption in Article 22 of the IT Framework Act. Superior Court of Québec, A.B. v. Google LLC, 2023 QCCS 1167, https://www.canlii.org/en/qc/qccs/doc/2023/2023qccs1167/2023qccs1167.pdf. As noted by judges: "Article 19.17.2 CUSMA does not require Canada to have an immunity provision that is identical to the expansiveness of the American provision, section 230(c)(1) CDA." *Id.* ¶ 182.

the European solution rests in focusing on a one-size-fits-all compromise to rule the legal system of each of its Member States instead of searching for compromises in areas of unharmonized domestic law. Thus, conditional immunity was held as a single standard to which liability in all areas of law in the Union must converge. Section 4 of the E-Commerce Directive greatly simplified the immunity part of § 512 of the DMCA by stripping it of its tricky parts.[11] This allowed technology companies to retain the benefits of the European Union's E-Commerce Directive regime by simply complying with more demanding U.S. copyright law. In practice, the much more detailed DMCA rules about notice-and-takedown choreography became the de facto standard across the world.[12]

The last two decades have largely validated the DMCA's conditional immunity as a feasible baseline approach to the regulation of internet communications that power global exchanges of ideas, goods, and services. However, the conditional immunity model has its limits. It was not designed to offer a complex solution for new challenges. Firstly, many of them were not known or debated at the time. Second, only a tiny fraction of humanity used the internet, and if people did use it, it was not a large part of their lives. At the time of the E-Commerce Directive's adoption in 2000, less than seven percent of the world's population used the internet.[13]

By 2016, a new mainstream sentiment concerning digital services started spreading in Europe and the United States. The Court of Justice of the European Union's (CJEU) newly invented "right to be forgotten" was rapidly taking off and putting pressure on the responsibility of search engines to individuals.[14] Facebook's neglect of content moderation in Myanmar exposed the grave risks of providers' chronic under-investment in less lucrative

11. E-Commerce Directive, O.J. (L 178), 1–16. The E-Commerce Directive did not incorporate general requirements, such as the implementation of a reasonable repeat-infringer policy (§ 512(i)(1)(A)), standard technical measures (§ 512(i)(1)(B)), or special requirements, such as lack of "a financial benefit directly attributable to the infringing activity" (§ 512(c)(1)(B)). On the other hand, in contrast to the DMCA, the ECD opens the doors much more extensively to injunctions.

12. Seng, *supra* note 9; Jennifer M. Urban, Joe Karaganis & Brianna Schofield, *Notice and Takedown in Everyday Practice*, UC BERKELEY PUB. L. RSCH. PAPER NO. 2755628 (2017), https://ssrn.com/abstract=2755628 ("Beyond its influence as a model, the DMCA also operates as de facto international law because the vast majority of notices are sent to US-based companies, which operate under it.").

13. *Individuals Using the Internet*, WORLD BANK, https://data.worldbank.org/indicator/IT.NET.USER.ZS (last visited Sept. 8, 2023).

14. Case C-131/12, Google Spain SL v. Agencia Espanola de Proteccion de Datos, ECLI:EU:C:2013:424 (Jun. 25, 2013).

markets.[15] The run-up to the 2016 U.S. elections inevitably politicized the topic of content moderation on social media. Social media in Europe was caught in the middle of the European migration crisis, which surfaced incredible amounts of organized support—but also toxic hate speech—among the general population.[16] It's likely that at this point, European governments began to question if self-regulation was the right approach. It became evident that the space that the conditional immunity model left to providers must soon be filled by regulation.

The DSA is the first comprehensive attempt to create a second generation of rules for digital services that rely on user-generated content. Unlike previous sectorial initiatives,[17] its approach is sweepingly horizontal. The DSA requires some level participation from both state and non-state institutions for its system of checks and balances to work, and some of its solutions can be "too European." However, the *principles* behind the DSA could be useful in other jurisdictions—perhaps even in the United States. The United Kingdom, which is currently developing its own set of post-Brexit rules, continues to build on some of the same principles as the DSA.

My hope is that these high-level principles might form the basis for a dialogue between liberal democracies about how to best regulate user-generated content services.[18] After all, if Europeans in the late 1990s could simplify and scale up the U.S. rules to fit their goals, maybe today other countries can do the same with the new E.U. rules. Having interoperable policies continues to be important for the flourishing of a truly global network of communications that generates unprecedented benefits for humanity.

---

15. Steve Stecklow, *Hatebook*, REUTERS (Aug. 15, 2018), https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/.

16. EUR. COMM'N, RACISM AND DISCRIMINATION IN THE CONTEXT OF MIGRATION IN EUROPE (Mar. 31, 2017), https://ec.europa.eu/migrant-integration/library-document/racism-and-discrimination-context-migration-europe_en; EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS, CURRENT MIGRATION SITUATION IN THE EU: HATE CRIME, (Nov. 2016), https://fra.europa.eu/sites/default/files/fra_uploads/fra-2016-november-monthly-focus-hate-crime_en.pdf.

17. The German and French parliaments previously adopted anti-hate speech rules that mostly imposed tight reaction periods for providers, see Netzwerkdurchsetzungsgesetz [NetzDG] [The Network Enforcement Act of 2017], Jan. 9, 2017, Bundesgesetzblatt, Teil I [BGBL I] at 3352 (Ger.); Loi 2020-766 du 24 juin 2020 visant a lutter contre les contenus haineux sur internet [Law 2020-766 of 24 June 2020 to Combat Hate Content on the Internet] [Loi Avia], Journal Officiel de la Republique Francaise [J.O.] [Official Gazette of France] (Jun. 25, 2008), p. 156.

18. For a broader debate, see MARTIN HUSOVEC, PRINCIPLES OF THE DIGITAL SERVICES ACT (Oxford Univ. Press forthcoming 2024).

II. FROM DMCA TO DSA

The European regulation of user-generated content services is clearly inspired by U.S. law. In this section, I first briefly explain how this has happened and then why, despite today's controversies, conditional immunity is an approach that has been arguably validated over the last two decades.

A. A BRIEF HISTORY OF LIABILITY EXEMPTIONS

Unlike the first liability exemption of its kind, § 230 of the CDA, which did not attract much stakeholder attention at the time,[19] § 512 of the DMCA is a product of hard negotiations between content industries and technology companies.[20]

The debate about the copyright liability of providers was power-charged by the 1995 White Paper issued by the Clinton administration's Information Infrastructure Task Force, which supported its view with two earlier rulings from U.S. courts regarding bulletin boards.[21] The White Paper presented strict direct copyright liability of providers, including internet access providers, as a given and argued that it would be "premature to reduce the liability of any type of service provider[.]"[22] The report implicitly encouraged plaintiffs to test the waters against all providers, not just bulletin boards. In 1995, the Church of Scientology sued another bulletin board operator, along with an internet access provider, Netcom, in a U.S. district court.[23] While the court quickly ruled that companies are not directly and strictly liable, it established that contributory knowledge-based liability remains an option.[24] The *Netcom* case undoubtedly put telecommunications companies, an established industry, on alert about

19. JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET 67 (Cornell Univ. Press 2019) ("Despite its monumental statements about a new, hands-off approach to the internet, the bill was virtually unopposed on Capitol Hill. Lobbyists focused primarily on the telecommunications bill's impacts on phone and cable television service.").

20. UNITED STATES COPYRIGHT OFFICE, SECTION 512 OF TITLE 17: A REPORT OF THE REGISTER OF COPYRIGHTS 18 (2020), https://www.copyright.gov/policy/section512/section-512-full-report.pdf.

21. Playboy Enters., Inc. v. Frena, 839 F. Supp. 1552 (M.D. Fla. 1993); Sega Enters. Ltd. v. MAPHIA, 857 F. Supp. 679 (N.D. Cal. 1994).

22. INFO. INFRASTRUCTURE TASK FORCE, INTELLECTUAL PROPERTY AND THE NATIONAL INFORMATION INFRASTRUCTURE: THE REPORT OF THE WORKING GROUP ON INTELLECTUAL PROPERTY RIGHTS 128 (1995), https://www.eff.org/files/filenode/DMCA/ntia_dmca_white_paper.pdf [hereinafter *White Paper*].

23. Religious Tech. Ctr. v. Netcom On-Line Comm. Servs., Inc., 907 F. Supp. 1361 (N.D. Cal. 1995).

24. Providers were not acting volitionally with respect to copyright-relevant acts, and thus cannot be held strictly liable. However, given that Netcom was served with notice, this triggered a duty to investigate the matter to avoid contributory copyright liability.

potential liability risks even though the outcome was favorable to them.[25] Those companies eventually lobbied to codify *Netcom* in the DMCA.[26]

After the White Paper's proposals failed in the 104th United States Congress,[27] the next Congressional session starting in January 1997 hoped to find a quick solution between opposing interests to successfully implement the World Intellectual Property Organization (WIPO) Internet Treaties.[28] In the legislative process, liability exemptions became a precondition to the passage of the entire piece of legislation.[29] As noted by the Senate Judiciary Committee Report, although the issue "[was] not expressly addressed in the actual provisions of the WIPO treaties, the Committee is sympathetic to the desire of . . . service providers to see the law clarified in this area."[30] It was understood that "without clarification of their liability, service providers may hesitate to make the necessary investment in the expansion of the speed and capacity of the internet."[31]

The Judiciary Committee report initially only included a liability exemption for mere conduits.[32] The final compromise with four liability exemptions—conduits, caching, hosting, and information location tools—only materialized after three months of direct negotiations between providers and content

25. JESSICA D. LITMAN, DIGITAL COPYRIGHT 128 (Prometheus Books 2d ed. 2006).

26. *See* H.R. Rep. No. 105-551, pt. 1 at 11 (1998), http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=105_cong_reports&docid=f:hr551p1.105.pdf ("As to direct infringement, liability is ruled out for passive, automatic acts engaged in through a technological process initiated by another. Thus, the bill essentially codifies the result in the leading and most thoughtful judicial decision to date: Religious Tech. Ctr. v. Netcom On-Line Commc'n Servs., Inc., 907 F. Supp. 1361 (N.D. Cal. 1995). In doing so, it overrules those aspects of Playboy Enters., Inc. v. Frena, 839 F. Supp. 1552 (M.D. Fla. 1993), insofar as that case suggests that such acts by service providers could constitute direct infringement, and provides certainty that *Netcom* and its progeny, so far only a few district court cases, will be the law of the land").

27. JESSICA D. LITMAN, DIGITAL COPYRIGHT 122 (Prometheus Books 2d ed. 2006).

28. *Id.* at 126, 130 ("After the bruising copyright fight in the last Congress, it wanted to satisfy the Hollywood and Silicon Valley communities but did not want to have to expend significant pollical capital to do so.").

29. *Id.* at 134–35.

30. S. Rep. No. 105-190, at 19 (1998). WCT only indirectly mentions the position of providers that can be found in an agreed statement to Article 8 which was the result of lobbying by providers and telecommunications companies who failed to include liability exemptions into the WIPO Internet Treaties themselves. *See* MIHALY FICSOR, THE LAW OF COPYRIGHT AND THE INTERNET: THE 1996 WIPO TREATIES, THEIR INTERPRETATION AND IMPLEMENTATION 509 (Oxford Univ. Press 2002).

31. S. Rep. No. 105-190 (1998).

32. H.R. Rep. No. 105-551 (1998).

owners.[33] The compromise text was already captured in the Commerce Committee in June 1998,[34] which argued that:[35]

> Title II preserves strong incentives for service providers and copyright owners to cooperate to detect and deal with copyright infringements that take place in the digital networked environment. At the same time, it provides greater certainty to service providers concerning their legal exposure for infringements that may occur in the course of their activities.

The DMCA was signed into law in October 1998. Congress was "keenly aware that other countries will use U.S. legislation as a model."[36]

In Europe, the European Commission published a communication to the European Parliament and Council in October 1996 explaining that providers will need legal assurances to be able to properly operate in the online market. The communication stated that:[37]

> Internet access providers and host service providers play a key role in giving users access to Internet content. It should not however be forgotten that the prime responsibility for content lies with authors and content providers. It is therefore essential to identify accurately the chain of responsibilities in order to place the liability for illegal content on those who create it . . . . The law may need to be changed or clarified to assist access providers and host service providers, whose primary business is to provide a service to customers, to steer a path between accusations of censorship and exposure to liability.

Two years later, the European Commission introduced the proposed E-Commerce Directive. Its Section 4 included three liability exemptions—conduits, caching, and hosting. At the time, only two European countries had liability exemptions. Germany adopted its two horizontal liability exemptions in July 1997[38] (termed the IuKDG) and Sweden adopted a law on bulletin

33. S. Rep. No. 105-190 at 7 ("These negotiations continued under the supervision of the Chairman for three months, from January to April, 1998."). *See* JESSICA D. LITMAN, DIGITAL COPYRIGHT 135 (Prometheus Books 2d ed. 2006).

34. H.R. Rep. No. 105-551 (1998).

35. *Id.*

36. S. Rep. No. 105-190 (1998).

37. Communication from the Commission on Illegal and Harmful Content on the Internet, COM (1996) 487 final, at 12–13 (Oct. 16, 1996).

38. Informations- und Kommunikationsdienste-Gesetz [IuKDG] [Act on Information and Communication Services of 1997] (Jun. 13, 1997), BGBL I at 52. Section 5 of IuKDG was elegantly condensed in the following four parts establishing the following: (1) liability is for own content remains to be governed by generally applicable law; (2) liability for other people's content on services that can be "used by others" ("die sie zur Nutzung bereithalten")

boards in May 1998.[39] Both the new German laws and DMCA made the basic distinction between services giving "access" and "space" to other people's information. Thus, unlike § 230 of the CDA, both the IuKDG and the DMCA differentiated liability exemptions based on the proximity of providers to users' actions. Conduits as distant facilitators were given the broadest immunity, while nearer hosts were granted more cautious exemptions based on their knowledge. In terms of scope, the German laws seemed more far-reaching, as they extended to conduits and all services which were being "made available for use[.]"[40] In contrast, § 512 focused on specific technical functions—conduits, caching, storage, and information location tools.

The main inspirations for Section 4 of the E-Commerce Directive were § 512 of the DMCA and Section 5 of the IuKDG. The Commission borrowed three liability exemptions from the DMCA, and a horizontal approach from the IuKDG. Unlike the U.S. copyright statute, the E.U. proposal was not driven by the need to implement the WIPO Internet Treaties but rather the European Union's desire to create an internal market without frontiers in the early stage of the internet's development. The newly found U.S. copyright compromise concerning the internet was thus extended to all areas of law.

The European Commission's proposal was adopted in June 2000. The Commission's approach followed the American definitions of categories of services and thus arguably narrowed down the scope of services which could rely on conditional immunity. For instance, the German provision could have easily covered information location tools, which were not given any explicit immunity.[41] In 2002, the E-Commerce Directive became law for fifteen E.U. Member States and, two years later, for another ten newly joined member states. As of now, both the E-Commerce Directive and the Digital Services Act apply across 27 member states. The Digital Services Act, as an E.U. regulation, is applicable directly without a need for local implementation. Post-Brexit, the United Kingdom so far has not repealed its implementation of the ECD liability exemptions, and E.U. case law until the end of 2020 continues

is possible only once they acquire knowledge; (3) liability for giving access to other's people content is barred; (4) blocking remains possible in accordance with generally applicable law.

39. Lag om ansvar for elektroniska anslagstavlor (Svensk forfattningssamling [SFS] 1998:112) (Swed.).

40. Section 5(2) of the IuKDG ("die sie zur Nutzung bereithalten").

41. Their qualification under hosting is complicated in the European Union due to questions about whether the information is "provided by" the indexed websites in all cases.

to be binding in British courts.[42] The United Kingdom is currently developing its own set of online safety rules that will supplement the existing exemptions.[43]

While the differences between the statutory language of the E.U. and U.S. laws were not insignificant, they were mostly reconcilable.[44] Generally, one can say the European Union simplified the DMCA—but also omitted some of its key components. In particular, the European Union omitted a liability exemption for information location tools and the DMCA's elaborate conditions for injunctive relief; the latter omission became a major point of divergence. Under E.U. law, injunctions were, in principle, left unconstrained if they conformed to the notion of "specific" monitoring.[45] The DMCA, in contrast, limited injunctions with a myriad of conditions.[46] As a result, while under § 512 of the DMCA all preventive injunctions—such as those imposing filters or website blocking—remained practically impossible, under Section 4 of the ECD they soon became the primary driver of European litigation efforts.[47] Eventually, the CJEU allowed plaintiffs who successfully litigated their grievances to seek injunctions that saddled providers with more responsibility to identify infringing content.[48]

The introduction of liability exemptions in the United States and European Union was clearly driven by the same rationale: to encourage investment by giving more legal certainty. As a result, the legal system can "steer a path

42. *See* The Electronic Commerce (EC Directive) Regulations 2002, SI 2001/2555 (Eng.), https://www.legislation.gov.uk/uksi/2002/2013, along with the European Union (Withdrawal) Act 2018, c.16, § 6 (UK), https://www.legislation.gov.uk/ukpga/2018/16/section/6/enacted.

43. *See* Online Safety Bill 2022-3, HL Bill [362] (UK), https://bills.parliament.uk/bills/3137.

44. Miquel Peguera, *The DMCA Safe Harbors and Their European Counterparts: A Comparative Analysis of Some Common Problems*, 32 COLUM. J. L. & ARTS 481, 481–82 (2009).

45. *See* E-Commerce Directive, art. 15(1), O.J. (L 178), 13; *id.*, r. 47, 6 ("Member States are prevented from imposing a monitoring obligation on service providers only with respect to obligations of a general nature; this does not concern monitoring obligations in a specific case and, in particular, does not affect orders by national authorities in accordance with national legislation.").

46. 17 U.S.C. § 512(j) (significantly limiting forms of injunctions) and 17 U.S.C. § 512(m) ("Nothing in this section shall be construed to condition the applicability of [liability exemptions on] a service provider monitoring its service or affirmatively seeking facts indicating infringing activity, except to the extent consistent with a standard technical measure ").

47. *See generally* MARTIN HUSOVEC, INJUNCTIONS AGAINST INTERMEDIARIES IN THE EUROPEAN UNION: ACCOUNTABLE BUT NOT LIABLE? (2017).

48. The biggest shift was brought by the CJEU in Case C-18/18, Glawischnig-Piesczek v. Facebook Ir. Ltd., ECLI:EU:C:2019:458 (June 4, 2019).

between accusations of censorship and exposure to liability."[49] The problem was acute to different degrees in different areas of law; however at the time, when CEOs of some technology companies were sentenced in criminal proceedings for distributing pornography, the concerns certainly were not trivial or overblown.[50] The growing national case law in the E.U. was seen as both too unpredictable and too unwieldy to provide clarity on how to reliably build a legal framework for the new environment that showed so much promise. Since user-generated content is so central to the digital *communications* network, the liability question was *the* question of internet regulation.

B. LIABILITY EXEMPTIONS AND SPECIFICITY OF THE INTERNET

The law has a key role in guaranteeing the shape and form of the internet. The decentralized nature of the internet as a network is inseparable from the underlying liability regime for those who facilitate its functioning. Without the sympathy of the law, there is no internet as we know it. In a hypothetical world where technology facilitates decentralization but the law provides incentives against it, no rational actors would have created spaces or tools without editorial control. A liability regime for the actions of others is a key incentive factor. Unless legislatures want to reinstate editors, some form of conditional immunity is necessary.

The European plan for most of the user-generated content services that host content is to ask victims to use nonjudicial notice-and-takedown systems and rely on the help of authorities, including courts, where possible. This mix of routes, while more generous to victims than the immunity-based framework of § 230 of the CDA, constrains victims' and the state's abilities to solve any social problem. But it does so for a good reason: to maintain the benefits of a decentralized communication network. By observing liability exemptions, everyone commits to constraining themselves in order to facilitate the emergence of an environment from which everyone can benefit. This is the essence of the digital social contract.

Strict liability, in contrast, demands total control, and such legal rules would become very expensive for society. By way of analogy, printers who are strictly liable for everything they print for others would inevitably need to first read and vet everything they print. Printing would become very slow and

49. Communication from the Commission on Illegal and Harmful Content on the Internet, COM (1996) 487 final, at 13 (Oct. 16, 1996).

50. In Germany, the law was also a reaction to the controversial CompuServe case. *See* Stefan Engel-Flechsig, Frithiof Maennel & Alexander Tettenborn, *Das neue Informations- und Kommunikationsdienste-Gesetz,* NJW 1997 2981, 2984 (1997).

expensive as a result, and people would be increasingly unable to use it to share ideas.

The link between such liability and freedom of speech has been recognized by the United States Supreme Court, the European Court of Human Rights (ECtHR), and the Court of Justice of the European Union in their human rights jurisprudence.[51] These highest courts set the limits for how user-generated services can be regulated by legislatures responsible for a little over 1 billion people.[52] At the moment, the strict liability of providers for user-generated content is treated on both sides of the Atlantic as unthinkable and fundamentally unconstitutional. U.S. and E.U. courts in unison continue to advocate for "medium-specific"[53] or "graduated and differentiated"[54] regulation that differs from regulation of editorial media. The European Court of Human Rights, for instance, despite its complex case law,[55] makes it clear

51. *See* Reno v. Am. C.L. Union, 521 U.S. 844 (1997); Case C-401/19, Poland v. Council & Eur. Parliament, ECLI:EU:C:2021:613 (July 15, 2021); MTE and Index.hu v. Hungary, App. No. 22947/13 (Feb. 2, 2016), https://hudoc.echr.coe.int/fre?i=001-160314.

52. To be precise: 690 million in the Council of Europe, of which 447 million are in the European Union, and then 331 million in the United States. *COE—Council of Europe 2023*, COUNTRY ECON., https://countryeconomy.com/countries/groups/council-europe (last visited Sept. 9, 2023) (noting that Russia is not a member anymore).

53. The "medium-specific" approach is relied upon by Judge Dalzell in Am. C.L. Union v. Reno, 929 F. Supp. 824, 873 (E.D. Penn. 1996) ("My examination of the special characteristics of internet communication, and review of the Supreme Court's medium-specific First Amendment jurisprudence, lead me to conclude that the internet deserves the broadest possible protection from government-imposed, content-based regulation.").

54. *See* Council of Eur., Recommendation on a New Notion of Media, CM/Rec (2011)7 ¶ 7 (2013), https://edoc.coe.int/en/media/8019-recommendation-cmrec20117-on-a-new-notion-of-media.html ("A differentiated and graduated approach requires that each actor whose services are identified as media or as an intermediary or auxiliary activity benefit from both the appropriate form (differentiated) and the appropriate level (graduated) of protection and that responsibility also be delimited in conformity with Article 10 of the European Convention on Human Rights and other relevant standards developed by the Council of Europe."), cited by the ECtHR in Delfi AS v. Estonia, App. No. 64569/09, ¶ 113 (June 16, 2015), https://hudoc.echr.coe.int/app/conversion/pdf/?library=ECHR&id=001-155105&filename=001-155105.pdf.

55. The European Court of Human Rights signaled that the Member of the Council of Europe might be exceptionally allowed to legislate that discussion forum providers should do more than only operate notice-and-takedown to avoid civil liability for hate speech. *See Delfi AS*, App. No. 64569/09. The decision is often mischaracterized as imposing a particular liability framework on the states. The case law only gives discretion to states to do this. Even more controversially, in a case concerning Facebook page administrators, the ECtHR also allowed the criminal financial liability of politicians for comments posted by others if they have some—albeit not specific—knowledge about those comments. Sanchez v. France, App. No. 45581/15, ¶ 162 (Sept. 2, 2021), https://hudoc.echr.coe.int/fre?i=001-211599. However, neither of the two rulings allows unconditional strict liability.

that "the notice-and-take-down-system could function in many cases as an appropriate tool for balancing the rights and interests of all those involved."[56]

The different treatment of the internet as a medium is not an act arising from a rose-tinted, naïve love for new technology.[57] It comes down to what the American Judge Dalzell in 1996 called "the special attributes of internet communication" that make it "the most participatory form of mass speech yet developed[.]"[58] The Court of Justice of the European Union referred to the internet as "one of the principal means by which individuals exercise their right to freedom of expression and information[,]"[59] and supported the view of the European Court of Human Rights that "user-generated expressive activity" is "an unprecedented platform for the exercise of freedom of expression."[60]

For Judge Dalzell and his colleagues in the late 90s, these "special attributes" were very low barriers to entry for speakers and readers leading to "astoundingly diverse content" and "significant access to all who wish to speak in the medium[.]"[61] For top European judges looking at it in the early 2010s, the special attributes of the internet are: its "accessibility"; its "capacity to store and communicate vast amounts of information"; its ability to support "user-generated expressive activity"; and its role in "facilitating the dissemination of information in general[.]"[62]

---

56. MTE and Index.hu v. Hungary, App. No. 22947/13, ¶ 91 (Feb. 2, 2016), https://hudoc.echr.coe.int/fre?i=001-160314 (presented as an application of the Grand Chamber decision in Delfi AS v. Estonia).

57. Discussing "internet exceptionalism" is beyond the space limitations of this Article, but the two essays worth reading on this are Mark Tushnet, *Internet Exceptionalism: An Overview from General Constitutional Law*, 56 WM. & MARY L. REV. 1637 (2015), and Tim Wu, *Is Internet Exceptionalism Dead?*, *in* THE NEXT DIGITAL DECADE: ESSAYS ON THE FUTURE OF THE INTERNET (Berin Szoka et al. eds., 2011), https://scholarship.law.columbia.edu/faculty_scholarship/1676.

58. Am. C.L. Union v. Reno, 929 F. Supp. 824, 867, 883 (E.D. Pa. 1996).

59. Case C-401/19, Poland v. Council and European Parliament, ECLI:EU:C:2021:613, ¶ 46 (July 15, 2021).

60. Delfi AS v. Estonia, App. No. 64569/09, ¶ 110 (June 16, 2015), https://hudoc.echr.coe.int/app/conversion/pdf/?library=ECHR&id=001-155105&filename=001-155105.pdf.

61. *Am. C.L. Union*, 929 F. Supp. at 877.

62. Case C-401/19, Poland v. Council and European Parliament, ECLI:EU:C:2021:613 (July 15, 2021), at ¶ 46 ("In the light of their accessibility and their capacity to store and communicate vast amounts of information, internet sites, and in particular online content-sharing platforms, play an important role in enhancing the public's access to news and facilitating the dissemination of information in general, with user-generated expressive activity on the internet providing an unprecedented platform for the exercise of freedom of expression"). The Grand Chamber is citing the ECtHR decisions in Cengiz and Others v. Turkey, Apps. No. 48226/10 and 14027/11, ¶ 52 (Dec. 1, 2015), https://hudoc.echr.coe.int/app/conversion/docx/pdf?library=ECHR&id=001-159188&filename=CASE%20OF%20

Any exemptions are naturally suspect to legislative favoritism towards the industry. And liability exemptions can be naturally fashioned in different ways. However, it has long been recognized that the DMCA-modelled liability exemptions are not necessarily major liability carve-outs when compared to ordinary applications of liability.[63] As noted by Advocate General Jääskinen, "these provisions are better qualified as restatements or clarifications of existing law than exceptions thereto."[64] This is also clear when looking at the text of § 512 of the DMCA, which incorporates many requirements of American copyright secondary liability.[65] Thus, while conditional immunities like those laid out in the ECD and DMCA might bring about some changes, they are usually not major liability carve-outs. The case for internet exceptionalism is somewhat stronger with the prohibition of general monitoring. However, its strongest legitimacy is in the protection against indiscriminate surveillance of people and their content, not as a rule to protect providers against increased costs.[66]

One could object that liability exemptions are therefore not *necessary* because courts would have gradually arrived at the right solution after years of litigation by simply applying general laws. While it is impossible to prove this with a counterfactual, the early history of liability in many countries[67] and even numerous recent examples of inconsistent case law show that legislative clarity has a unique value. For instance, while the recent ECtHR case law on liability does not in principle allow the states to depart far from knowledge-based immunity for hosts, the Court is clearly incapable of fashioning a predictable

---

CENG%C4%B0Z%20AND%20OTHERS%20v.%20TURKEY.pdf&logEvent=False, and Kharitonov v. Russia, App. No. 10795/14, ¶ 33 (Jun. 23, 2020), https://hudoc.echr.coe.int/fre?i=002-12866.

63. This is obviously different in the case of § 230 of the CDA, which lifts the constitutionally compelled immunity required by the First Amendment. *See generally* Eric Goldman, *Why Section 230 Is Better Than the First Amendment*, 95 NOTRE DAME L. REV. 33 (2019).

64. Case C-324/09, L'Oreál v. eBay, ECLI:EU:C:2010:757, ¶ 136 (Dec. 9, 2010).

65. *Compare* Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417 (1984), *with* Metro-Goldwyn-Mayer Studios, Inc. v. Grokster, Ltd., 545 U.S. 913 (2005).

66. As rightly pointed out by one of the reviewers, especially when low-cost means cannot be imposed due to the prohibition, the argument about existence of material carve-outs from the general framework might be valid. However, in such cases, the different treatment is not a result of favouring companies but favouring the privacy and expression rights of their users.

67. The early controversial U.S. cases concerned defamation law. *See, e.g.*, Stratton Oakmont, Inc. v. Prodigy Servs. Co., 23 Media L. Rep. 1794 (N.Y. Sup. Ct. 1995). The early German cases, on the other hand, concerned child abuse images and protection of minors *See, e.g.*, Entscheidungen des Amtsgericht München in Strafsachen [Munich Local Court] Az. 8340 Ds 465 Js 173158/95 (May 28, 1998).

test and is constantly creating endless pockets of new, sub-case law.[68] It seems that judges trained to engage in granular balancing are less interested in devising bright-line rules. Had the E.U. statutory law not been as clear as it was, human rights law would have hardly offered predictability.

C. THE NEED FOR A SECOND GENERATION OF RULES

The last two decades have drawn contours indicating many societal challenges that require solutions, ranging from: the protection of children; problems with hate speech or terrorism; to subversive activities that attack the basis of our democratic systems. All these problems are exacerbated by the "special features" of the internet as a medium: its lack of editorial approval, low barriers of entry (including omnipresent zero cost of services), incredible speed and scale of distribution, its broad social and geographical inclusiveness, and resilience of communications. Regulators across the globe are thus rightly considering how to address these challenges.

Simply pointing to the existing digital social contract seems insufficient when the clear legislative goal of the liability exceptions was to lay down incomplete and unrestrictive rules that would allow the medium to flourish. The tendency of some stakeholders to see liability exemptions as a magical limit on any future regulation mischaracterizes their key contribution. The key contribution is not in stopping any new rules from being adopted but in keeping one set of sufficiently enabling rules on the books. In any federal system, federal liability exemptions help to coordinate national or state laws by preempting national- or state-level experimentation. This is the added benefit of such rules both in the United States and European Union. However, this does not mean that such rules must be carved in stone. In fact, the E.U. and U.S. experiences both show that the inability of federal legislatures to update federal rules can lead states to test their limits.[69]

---

68. The two leading Grand Chamber cases, Sanchez v. France and Delfi AS v. Estonia, are basically painted as exceptions in other cases like MTE and Index.hu v. Hungary. App. No. 45581/15 (Sept. 2, 2021), https://hudoc.echr.coe.int/fre?i=001-211599; Delfi AS v. Estonia, App. No. 64569/09 (June 16, 2015), https://hudoc.echr.coe.int/app/conversion/pdf/?library=ECHR&id=001-155105&filename=001-155105.pdf; Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v. Hungary, App. No. 22947/13 (Feb. 2, 2016), https://hudoc.echr.coe.int/fre?i=001-160314.

69. In Europe, the lack of early Union legislation led Germany and France to adopt their own hate speech laws for social media. *See* Case C-131/12, Google Spain SL v. Agencia Espanola de Proteccion de Datos, ECLI:EU:C:2013:424 (June 25, 2013). In the US, the lack of any federal regulation led to state laws in Florida and Texas. *See* S.B. 7072, 2021 Leg. (Fla.), https://www.flsenate.gov/Session/Bill/2021/7072/; H.B. 20, 2021 Leg., 87th Sess. (Tex.), https://capitol.texas.gov/BillLookup/History.aspx?LegSess=872&Bill=HB20.

Additionally, the harms and victims of various societal challenges come in different forms. Some harms are amplified by the design of services; others are caused by other people and only facilitated by lack of intervention. Some victims of such harms lack means, while some are well-resourced; some can use technology to uncover violations of their rights, while others cannot. Before the DSA, Section 4 of the ECD left all these concerns to self-regulation or national experimentation. However, to effectively regulate a global network, the regulatory action must be big enough for global companies to start paying attention to it. For instance, despite three decades of European data protection law, it took the GDPR—which was adopted in 2016—to fully bring the laws to everyone's attention.

Horizontal liability exemptions, such as the one found in Chapter 2 of the Digital Services Act (formerly Section 4 of the E-Commerce Directive) are about creating breathing space for speech and markets while allowing enforceability of the rights of victims, but they do not address specific challenges. The rules of the first generation—§ 230 of the CDA, § 512 of the DMCA, and Section 4 of the ECD—all suffer the same insufficiency. They excel at coordinating expectations to encourage investment but fail at offering tools to solve a wide range of societal problems that emerged along with the use of these services.

The European Union's Digital Services Act is one example of how to update the digital social contract without undermining the decentralized nature of the internet. The DSA re-affirms democratic legitimacy for the rules of conditional immunity, and even extends them on margin. Providers' liability for user-generated content thus mostly does not change.[70] What changes are regulatory expectations when companies make their decisions about other people's content or behavior, and, for some providers, what they need to think about when designing digital services. These companies are accountable to the public through regulation. However, such regulation is specifically designed for them. Instead of fitting user-generated services into ill-suited preexisting categories, they are given a regulatory category of their own based on their size and technical functions.

70. The DSA introduces a few changes to the text of the liability exemptions, which arguably expands them; especially the mere conduit liability exemption (Article 4) now applies to a broader set of infrastructure services; hosting exemption receives some minor additions (e.g., Article 6(3)), which arguably already follow from the pre-existing case law; the newly inserted Article 7 about own investigation arguably will have limited effect, and again builds upon the case law.

III. THE TWO PILLARS OF THE DSA

The Digital Services Act has two main pillars: (1) due process requirements for content moderation, and (2) risk management obligations for services. Content moderation is defined and regulated as the process of decision-making that emerges from providers' reliance on the liability exemptions, such as hosting. Risk management focuses on the system and product design of services and invites providers to consider the broader effects of their advertising infrastructure, recommendation algorithms, and other systems. The table below provides an overview of all the main DSA obligations.

Table 1

Two pillars of rules	Technical activity	Company or service size	Main types of due diligence obligations imposed by the DSA
Content moderation	• Conduit • Caching • Hosting	Companies of all sizes	• Contact points or legal agents • Clarity of terms and conditions
		Medium-size[71] and bigger companies	• Content moderation reports
	• Hosting	Companies of all sizes	• Notice submission rules • Justification of decisions • Crimes notification to authorities
	• Online platforms (a subset of hosting services)	Medium-size and bigger companies	• Prioritization of trusted flaggers • Measures against abuse • Internal appeal systems • External appeal systems • Transparency of advertising and recommender systems
Risk management	• Online platforms (a subset of hosting services)	Medium-size and bigger companies	• Protection of minors • Dark patterns • Know-Your-Client obligations of marketplaces
	• Very large online platforms (VLOPs) • Very large online search engines (VLOSEs)	45 million average active monthly users, regardless of the company's size or turnover	Upon designation by the European Commission: • Risk assessment and auditing of all design features of the product • Enhanced data access for researchers to study risks • Crisis response mechanism • Special advertising transparency • Choice on recommender systems • Internal compliance officers

A. CONTENT MODERATION

The DSA recognizes that content moderation decisions by private companies can have a large impact on people's livelihoods and their freedoms

71. Small enterprises are defined by a Council Recommendation as "an enterprise which employs fewer than 50 persons and whose annual turnover and/or annual balance sheet total does not exceed EUR 10 million." Commission Recommendation of 6 May 2003 Concerning the Definition of Micro, Small and Medium-sized Enterprises, 2003 O.J. (L 124) 36.

to share and receive information from others. The last decade has shown that companies are not always willing to invest sufficient resources into such decision-making, especially in smaller countries or markets. The E.U. legislature's solution is to regulate the *process* through which such content moderation decisions are made.

Content moderation rules must be clear and predictable (Article 14(1)), and decisions must be based on existing policies (Article 14(4)). A wide range of content moderation decisions is subject to an obligation of individual explanation (Article 17) and annual transparency reporting (Article 15). Each decision must be subject to free internal appeals (Article 20) and potentially external dispute resolution (Article 21). In addition, many expectations are purposefully vague. Notification systems must be "user-friendly" (Article 16(1)), decisions made "in a timely, diligent, nonarbitrary and objective manner" (Article 16(6)), and enforcement practices must pay "due regard to the rights and legitimate interests of all parties involved" (Article 14(4)).

The above provisions set the rules for the *process side* of content moderation. Underlying contractual rules about acceptable content or behavior remain to be set by companies. However, providers' rule-making space is indirectly constrained by the limits placed on the procedure. Due to the obligation to disclose rules upfront (Article 14(1)), companies cannot retroactively change their policies, or invent sanctions ex post facto that have no basis in their existing rules (Article 14(4)). Providers can continue to contractually constrain speech beyond illegality according to their preferences; however, they must apply the rules in a nonarbitrary and non-discriminatory manner. Any contractual policies will be interpreted by out-of-court dispute settlement bodies which cannot consider "secret rulebooks" of any kind.

This clearly shows that the DSA does not take away all the content moderation discretion from platforms. It generally does not limit what legal content can be prohibited by providers under their community guidelines—that is a power that providers retain. Thus, if providers do not like how out-of-court bodies read their rules, they can change them and make them clearer. But once they put the rules in black and white, they cannot claim a contrary meaning without actually changing them. The DSA limits only some grossly unfair policies (Article 14(4)) that would likely already struggle with other areas of explicit legal prohibitions, such as consumer law.

The goal of these procedural guarantees is to script the process of content moderation into a tighter choreography that better reflects the impact of content moderation decisions on individuals. The mix of very specific procedural rules and vague aspirational regulatory expectations is meant to provide the basis for standard-setting but also a north star for content

moderation decision-making. For individuals, the rules give them more credible due process rights which go well beyond the standard delivered by markets alone.

Scholars like Douek criticize regulatory due process expectations as an unnecessary "process theatre"[72] which does not solve the overall problems because it resembles "using a teaspoon to remove water from a sinking ship."[73] But is that the right framing? First, for the affected individuals, even a teaspoon of hope that their grievances can lead to proper resolution are good enough reasons to institute them. This rationale is hardly diminished by the fact that such personal disputes do not resolve the larger problems. Second, the DSA tries to use the personal dimension of disputes as a source of broader learning, something favored by Douek, and as a pressure to improve the overall quality of the processes.[74] Finally, for very large online platforms, content moderation is only one part of their overall risk management.

B. RISK MANAGEMENT

The DSA's second pillar concerns risk management, which comprises a set of rules that address how companies design their products and other behind-the-scenes processes. Unlike the United Kingdom's upcoming Online Safety Bill,[75] the DSA legislatively and explicitly doses responsibility by the size or impact of the services. Risk must be mitigated only by digital services known as online platforms; that is, services that distribute user-generated content to the public as their main feature.[76] Platforms operated by micro and small companies—those employing less than fifty employees or earning less than ten million euros annually[77]—have no risk management obligations. The intuition

72. Evelyn Douek, *Content Moderation as Systems Thinking*, 136 HARV. L. REV. 526, 577 (2022).

73. *Id.* at 606.

74. *See* Digital Services Act art. 21, 2022 O.J. (L 277) against the background of a lab experiment concerning ADR system as a solution to rational bias against over-blocking. Lenka Fiala & Martin Husovec, *Using Experimental Evidence to Improve Delegated Enforcement*, 71 INT'L REV. OF L. & ECON. (2022), https://www.sciencedirect.com/science/article/pii/S0144818822000357.

75. The UK's Online Safety Bill, *supra* note 43, is still in the legislative process. According to the recent impact assessment, out of 25 thousand forecasted regulated organizations, roughly 20 thousand are likely micro. Online Safety Bill 2022-3, Impact Assessment ¶ 109, (UK), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061265/Online_Safety_Bill_impact_assessment.pdf.

76. Digital Services Act art. 3(i), 2022 O.J. (L 277).

77. *See* Commission Recommendation of 6 May 2003 Concerning the Definition of Micro, Small and Medium-sized Enterprises, 2003 O.J. (L 124) 36–41, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32003H0361.

behind this is that with more power comes more responsibility. For risk management, platforms are divided into two groups:

- In the *lower* tier, mid-sized or bigger companies are subject to limited and prescriptive rules covering design practices.
- In the *upper* tier, online platforms or search engines which serve more than 45 million monthly active users in the European Union are subject to a more expansive and vaguer set of rules: general risk management.

The companies in the lower tier must mostly think about how their product design protects *minors*, and against *manipulative* and *aggressive* practices—also known as dark patterns. The companies in the upper tier must do the same, plus much more. Specific businesses are designated as quasi-public squares where many Europeans meet and exchange. By state designation, they are placed under special regulatory dialogue with the European Commission and national authorities (regulators) about *any relevant risks to democratic institutions and individuals*, including risks to people's freedoms and well-being. Given that these are interests that are hard to delineate, the scope is very broad.

In the first round in Spring 2023,[78] the following digital services were designated:

- *Social media*: Instagram, LinkedIn, Pinterest, Snapchat, TikTok, Twitter, Facebook, YouTube
- *Marketplaces*: Alibaba, AliExpress, Amazon Store, Booking.com, Google Shopping, Zalando
- *App stores*: Apple AppStore, Google Play
- *Other*: Google Maps, Wikipedia
- *Search engines*: Bing, Google Search.

The newly imposed risk management obligations are clearly meant to legislatively complement liability assurances with some societal responsibilities as to trust, safety, and fairness in these services.

Risk management is a result of two realizations. First, the importance of design to the health of any ecosystem. This point has been reinforced by

78. *See* the designations published in European Commission Press Release IP/23/2413, Digital Services Act: Commission Designates First Set of Very Large Online Platforms and Search Engines (Apr. 25, 2023), https://ec.europa.eu/commission/presscorner/detail/en/IP_23_2413. For an explanation of the DSA's scope, see Martin Husovec, *The DSA's Scope Briefly Explained* (2023), https://ssrn.com/abstract=4365029. At the moment two platforms, Zalando and Amazon, are seeking invalidation of their designations before the General Court. Case T-367/23, Amazon v. European Commission (July 5, 2023); Case T-348/23, Zalando v. European Commission (June 27, 2023).

Francis Haugen's Facebook revelations[79] that put the spotlight on how amplification encourages certain types of behavior. Second, constant information and resource asymmetry between authorities (regulators) and providers realistically dictate that providers have the primary responsibility to find new solutions. The DSA's obligations relate more to the process or systems put in place. However, as shown below, this is more easily stated than practiced. In recent years, some type of "systemic regulatory approach" has been advocated by many scholars;[80] however, the details of such proposals differ significantly.

A particularly influential concept was Lorna Wood's and William Perrin's proposal which inspired the United Kingdom's Online Safety Bill (OSB). The proposal argued for the safety "by design" approach described as follows:[81]

> The regulator should be given substantial freedom in its approach to remain relevant and flexible over time. *We suggest the regulator employ a harm reduction method similar to that used for reducing pollution: agree tests for harm, run the tests, the company responsible for harm invests to reduce the tested level, test again to see if investment has worked and repeat if necessary* . . . . The regulator would then work with the largest companies to ensure that they had measured harm effectively and published harm reduction strategies addressing the risks of harm identified and mitigating risks that have materialised.

The framing of their model, including its placement under the statutory "duty of care" umbrella,[82] requires redistribution of responsibility for individual harms. This in turn evokes supervision of recommendation systems and product design features that change user behavior, including what individual content is being posted by them. While Wood and Perrin insist that content regulation is not the result of their approach,[83] they also envisage regulators'

79. Statement of Frances Haugen: Hearing before the S. Sub-Comm. on Consumer Protection, 177th Cong. (2021) (statement of Frances Haugen, former Facebook employee and whistleblower), https://www.commerce.senate.gov/services/files/FC8A558E-824E-4914-BEDB-3A7B1190BD49.

80. For an overview of (mostly) U.S. scholarship, see Kate Klonick's response to Douek, *supra* note 72. Kate Klonick, *Of Systems Thinking and Straw Men*, 136 HARV. L. REV. 339, 347 (2023).

81. LORNA WOODS & WILLIAM PERRIN, ONLINE HARM REDUCTION—A STATUTORY DUTY OF CARE AND REGULATOR 7, 13 (2019) (emphasis added), https://d1ssu070pg2v9i.cloudfront.net/pex/carnegie_uk_trust/2019/04/08091652/Online-harm-reduction-a-statutory-duty-of-care-and-regulator.pdf.

82. *Id.* at 29–30.

83. *Id.* at 12 (discussing types of content).

ability to limit "harmful behavior"[84] prophylactically.[85] It thus hardly avoids addressing the substance of the environment—the underlying rules of engagement for users.

In contrast, Evelyn Douek's proposal equally centers around systems but in a very different way:[86]

> Instead of focusing on the downstream outcomes in individual cases, it focuses on the upstream choices about design and prioritization in content moderation that set the boundaries within which downstream paradigm cases can occur .... And in *focusing on procedural accountability rather than the pursuit of some substantive conception of an ideal speech environment*, it is more politically feasible and less constitutionally vulnerable.

Thus, her "substance-agnostic approach"[87] is much more limited because it allows companies to experiment with any (legal) content policies. However, it seems to be focused on regulation of amplification, which, as explained by Keller, is not always easily substance-agnostic either.[88]

In any risk management system, the relationship between substance and process is the most difficult one. First, any proposal that tries to tackle "risks" or overall "harmful behavior" cannot ignore that on user-generated content services, what users say or do remains a key risk factor. While user behavior can be encouraged by the design of services, in some form it will continue to exist irrespective of this encouragement; usually, the risks on such services result less frequently from purely non-human external factors.

Managing the risks of crowds *often* requires telling individuals how they must behave. If authorities subject the occurrence of selected *illegal* user expressions to some metrics, the legitimacy of such policy is straightforward. The legislatures already agreed that such behavior is illegal, and the authority is only trying to enforce compliance. However, if authorities subject the occurrence of some *legal* expressions to the same metrics, they can easily end up policing the bounds of what people can say—the content of their communications. Putting direct quotas on user expression, when taken to its logical conclusion, means telling some people what they cannot say.

---

84. *Id.* at 48. *See* Graham Smith, *Speech Is Not a Tripping Hazard—Response to the Online Harms White Paper*, CYBERLEAGLE (June 28, 2019), https://www.cyberleagle.com/2019/06/speech-is-not-tripping-hazard-response.html (discussing consequences).

85. *See generally* Woods & Perrin, *supra* note 81.

86. Douek, *supra* note 72, at 585 (emphasis added).

87. *Id.* at 606.

88. *See* Daphne Keller, *Amplification and Its Discontents: Why Regulating the Reach of Online Content Is Hard*, 1, J. FREE SPEECH L. 227 (2021).

Whether addressing "harm" or "risk," the key litmus test is *who* sets the boundaries for the content of communications. One approach gives such power to decide to authorities; others leave it to individuals, platforms and legislatures.

- The *full* risk management approach gives the broadest power to authorities to ask companies about how their service design influences what happens on the platforms. Authorities observe, compare, analyze, and ask for changes, including by imposing tailored standards or quotas of "problematic" user behavior, regardless of the behavior's legality. The mandate of authorities thus extends to lawful but awful content and permits them to become surrogate legislatures policing the boundaries of free expression.
- The *limited* risk management approach shares the concerns about system design that might encourage various risks but stops before giving the authorities (agencies) the power to rewrite what lawful individual behavior should be banned or suppressed by quotas. This approach recognizes that authorities do not have the legitimacy of parliaments. Parliaments should remain responsible for setting the goalposts of illegal content of communications. If a specific risk or harm is particularly damaging, parliaments can move the goalposts further.[89] As a result, the authorities limit their demands regarding legal content to solutions that preserve people's agency by giving them freedom of choice; such solutions mostly empower or re-design the users' choice architecture.

Arguably, the Digital Services Act adopts the limited risk management approach. The DSA does *not* explicitly address the problem of whether the European Commission can require providers to change their contractual standards of "lawful but harmful content" as part of the risk management strategies.[90] However, in the absence of any explicit legal mandate, any attempts by the Commission to suppress specific legal expressions would arguably violate the rule of law.[91] The United Kingdom's Online Safety Bill is

89. In the UK, the self-harm debate led to the empowerment obligation and a proposal to create a new offence of "encouraging or assisting serious self-harm." Online Safety Bill 2022-3 Amendments, HL Bill [87] (later 362), p. 1 (2022), https://bills.parliament.uk/publications/51205/documents/3437.

90. Digital Services Act art. 35(1)(b), 2022 O.J. (L 277) speaks of "adapting their terms and conditions and their enforcement." In my view, this does not necessarily mean prohibiting lawful behavior or content. It speaks to the clarity and predictability of rules.

91. The argument is that Digital Services Act art. 34, 2022 O.J. (L 277) on its own is not sufficient to fulfil the human rights requirements under the E.U. Charter to legitimize prohibitions of speech to be "prescribed by the law." *See* Charter of Fundamental Rights of

currently moving in the DSA's direction too,[92] although the original proposal could have led to a full risk management approach.[93] The Australian Online Safety Act of 2021, however, seems to go the farthest by allowing authorities to ask for the removal of lawful but awful content.[94]

The limited risk management approach can be best explained in an analogy with managing risks during public protests. Imagine a public assembly protesting immigration policies that gathers in the streets of a city. The role of providers can be analogized to the position of protest organizers.[95]

The DSA designates the largest services in the European Union as controlled public spaces; it tasks their designers—the providers—to analyze risks created by bringing crowds together and to intervene if needed. What is the role of the state and providers in such cases?

The state can impose safety measures on organizers and protesters to protect them from others and others from them; and to avoid hurting bodies, property, or businesses. Physical safety measures benefit freedom of expression because they make everyone more comfortable in expressing their views. To achieve this, the authorities can ask organizers to take various safety

---

the European Union art. 52, 2012 O.J. (C 326) 391. For an excellent article on the rule of law requirement in this context, see Graham Smith. *Online Harms and the Legality Principle*, CYBERLEAGLE (2020), https://www.cyberleagle.com/2020/06/online-harms-and-legality-principle.html ("[T]he regulator's views about harm would sit alongside, and effectively supplant, the existing, carefully crafted, set of laws governing the speech of individuals.").

92. Douek, *supra* note 72.

93. The UK government construed Wood and Perrin's proposal in its initial proposal of the Online Safety Bill by creating a controversial clause about safety duties for "harmful but lawful" content for adults. The relevant clause was dropped and the bill left with an empowerment obligation under the system known as "triple lock" in the later versions of the Bill. *See* Online Safety Bill 2022-3, HL Bill [362], § 12 (UK), https://bills.parliament.uk/bills/3137 (creating a duty for some services to "include in a service . . . features which adult users may use or apply if they wish to increase their control over content" that "reduce[s] the likelihood of the user encountering" or "alert" users to some types of content, such as hate speech, self-harm or eating disorders).

94. *See Online Safety Act 2021* (Cth), pt 4, 9 (Austl.), https://www.legislation.gov.au/Details/C2021A00076. According to Professor Nicolas Suzor, "The classification scheme has long been criticized because it captures a whole bunch of material that is perfectly legal to create, access and distribute." *See* Ariel Bogle, *Australia's Changing How it Regulates the Internet—and No-one's Paying Attention*, ABC NEWS (Sept. 20, 2022),https://www.abc.net.au/news/science/2022-09-21/internet-online-safety-act-industry-codes/101456902.

95. One of the reviewers made an excellent point, with which I nevertheless do not fully agree. The reviewer argues that a better analogy would be with the owner or operator of the property. In my view, this would evoke a very passive role of the platforms that do not influence the created risks by the design of their services. While the metaphor of organisers might better fit social media with active recommender systems than Wikipedia (also an online platform), my illustration is meant to show how self-imposed rules should not be adopted by authorities as the reason for intervention for otherwise legal protests.

measures particularly to prevent illegal behavior by protesters or counter-protesters, including proscribing the use of excessive disruption or noise. However, beyond illegal modes of expression, the *authorities* cannot control who speaks or protests, what posters or chants they use, or where they present them. That said, *organizers* can go beyond illegality, whatever their motivation. They can self-impose stricter rules on crowds.

Imagine now that this public assembly has two teams of rule enforcers dressed in red and blue jackets. Red enforcers represent the state, and they can only intervene when protesters violate a set of red rules—the behavior that the legislature has determined to be illegal. Blue enforcers are paid by organizers. They are the analogue of content moderators. Because organizers want a legitimate assembly where families can gather, they ask all the participants to respect some of their own basic rules. These blue rules differ from red rules. Among other things, they allow organizers much earlier intervention. For instance, they can say that posters with profanities are not permitted because they are likely to lead to illegal behavior.

For efficiency reasons, the state will expect blue enforcers to also enforce red rules. This is the analogue of delegated enforcement in which providers engage daily when they remove illegal content. However, red enforcers *cannot* enforce blue rules. Blue rules are the analogue of contractual self-restraint that platforms adopt to make their services appealing to users and advertisers. Red enforcers cannot turn a self-imposed ban on profanities against the organizers to end the protest or arrest protesters. To justify such intervention, authorities must stick to the red rules. Logically, they cannot tell organizers what blue rules to adopt either because that is the prerogative of legislatures. The state can, however, require that protesters inching closer to escalation must take extra measures to keep bystanders safe from violence.

The DSA's very large online platforms (VLOPs) and very large online search engines (VLOSEs) manage huge crowds constantly. As a result, they must periodically assess the risks, submit their reports to auditors, and follow up in case the auditors are not satisfied. The entire dossier of documents is then submitted to the European Commission for the ultimate assessment and release for the public to see and criticize.

IV. PRINCIPLES FOR A NEW GENERATION OF RULES

Now that the reader is familiar with the rules in the Digital Services Act, I would like to extract some of the main principles that define the regulatory approach. As pointed out by Daphne Keller, "differences between American and European approaches shouldn't prevent us from finding common ground

on other functional aspects of platform regulation."[96] The DSA has a lot to offer, but one needs to look beyond the exact wording and "under the hood" to understand the thinking. In my view, the following set of principles can be derived from the DSA and could serve as "common ground" to guide the legislative design of a new generation of rules:[97]

1. Accountability, not liability
2. Horizontality of regulations
3. Shared burden: everyone is responsible
4. Empowerment of users
5. Ecosystem solutions

A. ACCOUNTABILITY, NOT LIABILITY

Platforms as facilitators of user-generated content cannot be expected to bear the liability burden of conventional publishers, such as newspapers. As much as their content moderation might resemble quasi-editorial functions, the special features of the internet demand different legal regimes. The existence of some sensible legal immunities for liability generated by the actions of others is the basic precondition of the viability of the user-generated services which harness the internet's special benefits. These include no requirements for editorial approval, low barriers of entry, incredible speed and scale of distribution, broad social and geographical inclusiveness, and resilience of communications. Instead of devising restrictions which may negate these advantages, the focus should be on how to align providers' business operations with socially optimal practices that maximize freedoms of individuals—thus making the businesses more accountable to public interest.

Prior to the DSA, most of the relevant laws tried to influence providers' behavior by threatening them with accessory liability for what their users do.[98] Save for some areas of law, most notably intellectual property law in the European Union,[99] the courts often faced a binary decision: impose liability, with all its consequences, or deny it entirely and confirm a liability exemption. The DSA ends this binary. Self-standing regulatory expectations created by the

96. Daphne Keller, *For Platform Regulation Congress Should Use a European Cheat Sheet*, HILL (Jan. 15, 2021), https://thehill.com/opinion/technology/534411-for-platform-regulation-congress-should-use-a-european-cheat-sheet/.

97. *See* MARTIN HUSOVEC, PRINCIPLES OF THE DIGITAL SERVICES ACT (Oxford Univ. Press forthcoming 2024).

98. *See* Martin Husovec, *Remedies First, Liability Second: Or Why We Fail to Agree on Optimal Design of Intermediary Liability?*, *in* THE OXFORD HANDBOOK OF ONLINE INTERMEDIARY LIABILITY (Oxford Univ. Press 2020) (criticizing a one-size fits all approach).

99. *See* MARTIN HUSOVEC, INJUNCTIONS AGAINST INTERMEDIARIES IN THE EUROPEAN UNION: ACCOUNTABLE BUT NOT LIABLE? (Cambridge Univ. Press 2017).

legislature give courts and authorities a third option. A failure to satisfy such expectations is enforced separately. Thus, similar to banks that are usually not liable for the illegal financial transactions of their clients, they can still be held accountable and fined for not adopting the right anti-money laundering processes.

The DSA leaves existing liability exemptions almost intact. Section 4 of the ECD is incorporated into Chapter 2 of the DSA. Its novelty is in the creation of new regulatory expectations named "due diligence obligations" that are foreseen in Chapter 3. They are unrelated to legal immunities for third-party content. As noted by Recital 41 of the DSA, "[t]he due diligence obligations are independent from the question of liability of providers of intermediary services which need therefore to be assessed separately." If due diligence obligations are violated, they trigger a separate enforcement system envisaged by the DSA; they do not expose providers to a flood of claims for individual grievances. Due diligence obligations aim to improve the operations of systems and procedures that companies are using to moderate users' content or manage other overall risks.

To illustrate this, consider the following example. In American copyright law, under § 512(i) of the DMCA, a failure to terminate accounts of repeat infringers leads to the loss of a liability exemption and thus the potential joint liability of providers for the actions of users who infringe copyright when using their services. In European copyright law, such failure has no impact on liability exemptions. However, post-DSA, a failure to terminate accounts of repeat infringers can lead to a violation of Article 23(1) of the DSA, which can be enforced privately or publicly even though the liability exemption continues to apply. Thus, in both cases, the consequences are substantially different.

The DSA's accountability-but-not-liability design was not an automatic policy choice. In the legislative process, the European Parliament strongly pushed to make the liability exemptions dependent on compliance with due diligence obligations. Thus, any violation of Chapter III of the DSA would make liability exemptions unavailable. The opposite approach, where due diligence duties act as preconditions, exists under Section 79 of the Indian Information Technology Act (2000),[100] and is being proposed by Professor

100. Section 79 of the Indian Information Technology Act states that "the intermediary observ[ing] due diligence while discharging his duties under this Act and also observ[ing] such other guidelines as the Central Government may prescribe in this behalf." The Information Technology Act, 2000, § 79. Part II(4)(4) of the Indian Ministry Guidance, https://mib.gov.in/sites/default/files/IT%28Intermediary%20Guidelines%20and%20Digital%20Media%20Ethics%20Code%29%20Rules%2C%202021%20English.pdf (imposing filters on "significant social media intermediaries").

Danielle Citron as a solution for the revision of § 230 of the CDA in the United States.[101] Under such a system, the liability exemptions would become a truly hard-earned "prize" or a "privilege" given only to those who respected the DSA in its entirety. The more due diligence obligations are added to the list, the more impossible walking of the tightrope becomes. The E.U. legislature consciously decided against this approach—for good reasons, as it would basically nullify the existence of liability exemptions.

In the liability framework, the lack of diligence puts providers at risk of being an accessory to the entire wrongs of others. On the other hand, the accountability framework blames them only for not giving some specific assistance.[102] The legal culpability implied in the two settings is very different and it translates into the seriousness of the consequences for the platforms. While liable platforms face injunctions and joint liability for damages and are called to account by many victims who were wronged by the actions of others, accountable platforms only face the pain of enforcement efforts to bring them into compliance. Thus, while liable platforms restore a lawful state by making the victims whole, accountable platforms restore it by simply adjusting their behavior in ways that comply with regulatory expectations.[103]

If accountability is further narrowed down to *systemic* legal obligations in the design and operation of systems and processes, the difference is even more significant.[104] Under such systems, if a provider violates a systemic due diligence obligation, only one obligation to correct the outcome is owed to individuals or regulators. In contrast, if such obligation is embedded into a liability exemption, one failure to operate a specific policy leads to separate

101. Danielle Keats Citron, *How To Fix Section 230*, VA. PUB. L. & LEGAL THEORY RSCH. PAPER NO. 2022-18 (2022), https://ssrn.com/abstract=4054906 (arguing that § 230 should be narrowed in scope, and made subject to duties of care that can be further fleshed out by administrative agencies).

102. This should hold true for both public and private enforcement. Even for damages claims for violations of due diligence (Digital Services Act art. 54, 2022 O.J. (L 277)), the damage must be causally connected with the violation of the diligence obligation (Digital Services Act recital 122, 2022 O.J. (L 277) and only compensate the corresponding part of the damage. Thus, damages caused by third parties who uploaded the content are distinct.

103. Arguably, there are situations where liability exemptions will be lost, due diligence obligations violated, and the damage caused by a third party is closely related to that caused by violation of a due diligence obligation. In such cases, the DSA can indicate to national law that a component of the duty of care for domestic liability rules was violated. However, in many cases, the two harms are unlikely to be related (e.g., transparency rules or non-arbitrariness standards hardly relate to damage caused by third-party content).

104. The primary example of such obligations is Digital Services Act art. 21(2), 2022 O.J. (L 277) ("[E]ngage, in good faith, with the selected certified out-of-court dispute settlement body with a view to resolving the dispute"). Other examples are Article 22(2)(c), Article 23, and arguably many open-ended standards of Article 20(4).

debts to many who were wronged. Accountability for systemic obligations means owing one type of assistance to all affected people, while liability for others means owing all wronged people full liability for the actions of many other people. The difference is stark. Moreover, the DSA prohibits super-compensatory damages for due diligence violations (Article 54).

B. HORIZONTALITY OF REGULATIONS

The second principle implicit in the DSA's and ECD's design is its horizontal character. The horizontal approach cuts through the entire legal system and thus creates baseline expectations. Sectorial rules remain possible; however, they are forced to interact with the horizontal approach. In the European Union, the DSA thus becomes a *digital civil charter* that shines through the entire legal system and radiates minimum rights of individuals. Unless the European legislature suspends it in various areas, it creates a baseline that holds across the entire ecosystem of user-generated content services. In the DSA, the horizontality of liability exceptions is complemented by the horizontality of due diligence obligations. This supports my earlier argument about the updated digital social contract for user-generated content services. For instance, in the European system, the DSA's rules substantially improved the situation under several sectorial rules dealing with copyright issues.[105]

The horizontality of rules is not only useful for complying companies and individuals, but it also prevents gaming the system. The typical problem with sectorial rules imposing different standards is that they invite regulatory arbitrage. For instance, in the US, where Section 230 of the CDA provides even post-notification immunity to hosting services, there is a strong incentive to formulate any claims as copyright issues because § 512 of the DMCA is much more accommodating.[106] This turns defamation claims into copyright claims and distorts copyright policy in the long run. In a situation where the identity of claims is fluid and the plaintiffs can shop around for the strongest

105. The DSA updated the safeguards applicable under Article 17 of the Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC. *See* Martin Husovec, *Mandatory Filtering Does Not Always Violate Freedom of Expression: Important Lessons from Poland V. Council and European Parliament*, 60 COMMON MKT. L. REV. 173 (2023); João Pedro Quintais & Sebastian Felix Schwemer, *The Interplay between the Digital Services Act and Sector Regulation: How Special is Copyright?*, 13 EU. J. OF RISK REGUL. 191 (2022).

106. The most famous of the abuses of this kind are U.S. doctors asking for copyright assignment to text of future reviews to be able to require their takedown. *See* Mike Masnick, *Why Doctors Shouldn't Abuse Copyright Law to Stop Patient Reviews*, TECH DIRT (Apr. 14, 2011), https://www.techdirt.com/2011/04/14/why-doctors-shouldnt-abuse-copyright-law-to-stop-patient-reviews/.

cause of action, diverging standards for different legal areas are bound to cause regulatory arbitrage. The only way to avoid this is to adopt one set of uniform rules for all areas of law.

Horizontality also allows for better balancing of different trade-offs. For example, protecting minors might come at the expense of the freedoms of adults. Enforcement of hate speech policies can have unintended effects on legitimate discourse. Having a holistic policy allows the regulators to better balance one against the other, as their mandates extend to both. Thus, the European Commission, when looking at risks and technological solutions, must equally consider the under-detection of hate speech and over-blocking of legitimate speech. Given that content moderation and risk management stretch into all areas of human interactions, having the broadest possible focus is key to any balanced policy.

Politically, the horizontal approach also moderates the excessive strength of some interest groups because it broadens the conversation and dilutes their voice with the equally valid concerns of others. The E-Commerce Directive and the Digital Services Act could hardly have been adopted as sectorial measures. In fact, both the American and European examples show that copyright rules, an area that powerful lobbies of interest groups exercise influence over, constantly diverge from the baseline in favor of copyright holders. Section 512 of the DMCA is stricter than § 230 of the CDA. Similarly, Article 17 of the Copyright DSM Directive is stricter than Article 6 of the DSA.[107]

C. SHARED BURDEN: EVERYONE IS RESPONSIBLE

The DSA renews democratic support for the shared burden model for societal risks on digital services. Under the principle of shared burden, everyone is expected to play their part—to do something to protect oneself. It also means resisting the temptation to blame one actor for all ills.

In liability systems around the world,[108] it is an established principle that if victims contribute to their own damage by failing to exercise due care, the

107. Article 17 of the CDSM Directive introduces a system of strict liability for unlicensed content unless case providers can meet very strict cumulative conditions: inability to obtain a license, the stay-down obligation, and notice-and-takedown system. *See* Directive 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC Council Directive 2019/790, art. 17, 2019 O.J. (L 130) 92.

108. *See Principles of the European Tort Law (PETL), Art. 8.1.1*, http://www.egtl.org/PETLEnglish.html ("Liability can be excluded or reduced to such extent as is considered just having regard to the victim's contributory fault and to any other matters which would be relevant to establish or reduce liability of the victim if he were the tortfeasor."); Martin Turck,

person who is otherwise liable will face decreased or no liability. This principle of comparative negligence was famously formulated by Lord Ellenborough in 1809 who said that: "One person being in fault will not dispense with another's using ordinary care for himself."[109] Arguably, the ECD builds upon this principle in the design of its liability exemptions, and the DSA designs its due diligence obligations the same way.

Under liability exemptions, victims or their representatives must notify providers about infringing content or seek redress before authorities, and providers must act upon notifications or state-issued orders. Providers are usually not expected to prevent all individual grievances; instead, hosting providers must investigate them mostly once they are brought to their attention. Even the ex ante risk management due diligence obligations do not change that. Reporting illegal content, disputing providers' decisions, organizing with others, and learning and teaching others how to avoid risks remain the key ingredients of the DSA's content moderation system.

One of the expressions of the shared burden principle is also the prohibition of general monitoring in the ECD and DSA. Article 15 of the ECD, now Article 8 of the DSA, prohibits the following:

> No general obligation to monitor the information which providers of intermediary services transmit or store, nor actively to seek facts or circumstances indicating illegal activity shall be imposed on those providers.

The provision thus also embodies the idea[110] that the law generally structures the allocation of responsibilities to various actors. Providers are not subject to general obligations to intervene in other people's affairs. Such implicit allocation is not exhausted by the liability exemptions. This is also why any other rules imposed on providers, such as injunctions or any permitted national regulatory expectations, remain curtailed. As much as the burden under the liability exceptions system is shared, so must the burden under the accountability for risk management system be similarly split.

The sharing of the burden under liability exemptions allowed the user-generated content universe to flourish because it spreads responsibility and

Contribution Between Tortfeasors in American and German Law--A Comparative Study, 41 TUL. L. REV. 1 (1966–1967); Giuseppe Dari-Mattiacci & Eva S. Hendriks, *Relative Fault and Efficient Negligence: Comparative Negligence Explained*, 9 REV. OF LAW & ECON. 1 (2013).

109. Butterfield v. Forrester, Eng. Rep. 926, 927 (1809).

110. Advocate General Øe in his Opinion in C-401/19, ¶ 106 ("I am inclined to regard the prohibition laid down in Article 15 of Directive 2000/31 as a general principle of law governing the internet, in that it gives practical effect, in the digital environment, to the fundamental freedom of communication.").

thus expectations. Burden sharing under the accountability-for-risk-management framework will be equally crucial to avoid moral hazard.[111] While the DSA clearly puts accountability for risks on VLOPs/VLOSEs, it does not require the eradication of risks. Not all risks can be controlled by providers in the same way. While inherent risks cannot be mitigated at all, other risks can be increased by the behavior of providers, their users, or third parties.

For instance, the risk of fraud via digital scams depends not only on platforms' protective systems but also on their users' behavior, skills, and awareness. Providers can do a lot to prevent such scams; however, they can only partly influence users' behavior, skills, and awareness. The risk thus needs to be distributed, and users must share their part of the burden. This is how we deal with risks in most areas because protecting people against their own irresponsibility sometimes only breeds more irresponsible behavior.[112] The same starting point should be used to approach the regulation of issues such as the manipulation of votes by disinformation campaigns. The VLOPs' and VLOSEs' accountability for these harms is significant, but not absolute and not exclusive.

This brings me to my next principle.

D. USER EMPOWERMENT

The users can only be asked to learn how to share part of the risks if they are able, and thus empowered, to mitigate them. The principle of user empowerment means that ultimately, users can share only parts of those risks that they are given a chance to control. Typically, this means the provision of tools that grant people agency in deciding what they wish to see and from whom. If platforms leave little agency to users, they should assume more risks. The more agency users gain, the more they can control their own digital experience. Thus, undeniably, more user empowerment means less central responsibility of providers, which might not appeal to everyone. But it does not mean that such tools will allow providers to shrug off any accountability for risks; if coupled with reasonable expectations on the users' side, control given to users can at best reduce it.

The DSA tries to give users new levers of control over their user experience, such as the ability to challenge decisions, receive compensation for

111. *See generally* John M. Marshall, *Moral Hazard*, 66 AM. ECON. REV. 880 (1976).

112. For instance, in the EU, liability for unauthorized payments, such as those caused by phishing attacks, is primarily with banks. However, if clients behave grossly negligently, the banks do not have to compensate the clients. *See* Article 73 of the Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on Payment Services in the Internal Market, Amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC.

moderation mistakes, rely on representation before platforms, benefit from new parental tools and choice on recommender systems. As explained by Recital 40 of the DSA, the due diligence obligations:[113]

> should aim in particular to guarantee different public policy objectives such as the safety and trust of the recipients of the service, . . . the protection of relevant fundamental rights enshrined in the Charter, the meaningful accountability of those providers and the *empowerment* of recipients and other affected parties, whilst facilitating the necessary oversight by competent authorities.

Thus, empowerment of individuals is encoded in the DSA and invites providers to harness its power. The trade-off for VLOPs/VLOSEs is clear. Relinquish part of control in exchange for lesser accountability for risks or keep full control and assume more responsibility for what transpires on the platform. Risk-sharing is thus an incentive to delegate to users and enhance their agency as individuals with free will and preferences.

When I am talking about empowerment tools, I do not mean the obvious tools. Realistically, all platforms give users some agency in their digital experience. We all want to follow people based on our preferences and block people who cross our personal red lines.[114] However, platforms still assume too much central control over many decisions where the personal preferences of their users can legitimately diverge. By definition, this is most important for the category of legal content that can be controversial to host. While few users will diverge on their preferences for commercial spam, many might have different sensitivities for shocking, sensational, nude, or vulgar content.

In the literature, Fukuyama and others have argued for empowerment through a system of middle-ware tools that could help users to personalize their content moderation experience.[115] The idea of polycentric content moderation that puts users in charge of more decisions arguably already exists, however, before the DSA could not have been legally compelled. Consider a new start-up, TrollWall,[116] that offers social media page administrators a machine learning-based content moderation tool that is meant to address the slow removal of illegal content by Facebook, but also offers a scalable solution

113. Digital Services Act recital 40, 2022 O.J. (L 277) (emphasis added).

114. Naturally, any preference for illegal content is simply illegal and thus irrelevant.

115. FRANCIS FUKUYAMA, BARAK RICHMAN, ASHISH GOEL, ROBERTA R. KATZ, A. DOUGLAS MELAMED & MARIETJE SCHAAKE, REPORT OF THE WORKING GROUP ON PLATFORM SCALE, https://fsi9-prod.s3.us-west-1.amazonaws.com/s3fs-public/platform_scale_whitepaper_-cpc-pacs.pdf.

116. *See AI Autopilot for Comment Moderation*, TROLL WALL, https://www.trollwall.ai/ (last accessed Sept. 26, 2023).

to preserve the civility of online discussions. This tool gives administrators the ability to adjust content categories, sensitivity, and what should happen with the detected content. Although the tool is offered by a third party to page administrators,[117] Facebook has a key role in creating APIs that facilitate it and approves such apps for distribution in its platform. While far from being error-free, the tool gives administrators more agency to deal with problems with a scale that is prohibitively big for full human oversight. The DSA can pave the way to more of such tools that puts users and other individuals in charge.

E. ECOSYSTEM SOLUTIONS

If the responsibility for societal challenges is shared, everyone needs to be part of the solution. While providers and the state navigate their respective roles, civil society holds both to account.

Countering extremism or disinformation can be successful only if providers are assisted by an ecosystem of actors, such as trusted NGOs who notify the content, fact-checkers, journalists, or researchers. One of the shortcomings of the first generation of rules like the DMCA, CDA, and ECD is their preoccupation with providers and the little consideration they pay to those other players in the ecosystem.[118] Under the E-Commerce Directive, only platforms were relieved of liability. The other parties involved in solving the societal challenges in play were not given any specific tools to do their work. The self-regulatory approach was meant to solve this in the European Union. However, this often led to disparate arrangements across different services that can be taken away from civil society at the whim of new owners or leadership of providers.[119] For civil society, disparities mean difficulties in scaling the response.

The Digital Services Act puts the ecosystem front and center. It recognizes that content moderation is a product of decision-making by providers, but its quality is equally dependent on inputs (the quality of notifications) and feedback (the ability of users to correct the mistakes).

---

117. Naturally, Facebook can offer its own tools to page administrators, but these have so far very limited usefulness, especially in smaller markets. Thus, one can see how user empowerment can play out in small.

118. Jessica Litman has argued that the DMCA "sells the public short." And yet, § 512 DMCA at least includes some safeguards—even if ineffective in practice—such as details for notices (§ 512(c)(3)) and rules on counter-notice (§ 512(j)) or misrepresentation (§ 512(f)). *See* JESSICA D. LITMAN, DIGITAL COPYRIGHT 145 (Prometheus Books, 2d ed. 2006).

119. Brian Fung, *Academic Researchers Blast Twitter's Data Paywall as 'Outrageously Expensive,'* CNN (Apr. 5, 2023), https://edition.cnn.com/2023/04/05/tech/academic-researchers-blast-twitter-paywall/index.html.

On the side of inputs, the DSA tries to incentivize the quality of notifications. Providers are tasked with designing their submission interfaces in user-friendly ways to help other actors with their work.[120] It gives preferential treatment to trusted flaggers who have a track record of quality.[121] Trusted flaggers that abuse their position might be suspended or have their certification removed by regulators.[122] Providers are asked to suspend or terminate the accounts of those who repeatedly submit abusive notifications or manifestly illegal content.[123] The DSA encourages standardization[124] of how notices are exchanged which should lead to the emergence of more automated cross-platform solutions.

On the side of feedback, the DSA tries to decrease the information asymmetry between providers and their content creators. Providers must properly disclose their rules up front and describe what automated tools they use to enforce them.[125] They must issue individualized explanations for a wide range of content moderation decisions and allow appeals free of charge.[126] If content creators or notifiers are dissatisfied, they can file external appeals to out-of-court dispute resolution bodies.[127] The providers must pay for the complainant's costs of initiating external appeals whenever they lose cases, which should motivate them to improve the quality of their decisions internally.[128] Specialized organizations can be included in the dispute resolution process, thus allowing content creators to improve the quality of their representation.[129] Consumer groups are given a collective redress in the form of injunctions which can be sought to cure noncompliance.[130]

In the risk management pillar, the DSA asks researchers, civil society, and auditors to formulate relevant risks and invent new ways to mitigate them. For the largest digital services, regulators conduct a regulatory dialogue about societal challenges in public to intensify scrutiny.

120. Digital Services Act art. 16(1), 2022 O.J. (L 277),
121. *Id.* art. 22.
122. *Id.*
123. *Id.* art. 23.
124. *Id.* art. 44(1).
125. *Id.* arts. 14–15.
126. *Id.* arts. 17, 20.
127. *Id.* art. 21.
128. For an empirical test of this proposition, see *supra* note 74. Given that the system offers a more credible remedy, one can also expect that the use of it will increase, thus the impact will be higher than under the current system, where no independent third party is involved, and the only available remedy—courts—are not as de-risked for the complainants.
129. Digital Services Act art. 86, 2022 O.J. (L 277).
130. *Id.* art. 90.

In other words, the DSA gives other actors in the digital ecosystem tools that they can rely on when protecting private or public interests. By doing this, the DSA heavily relies on societal structures that the law can naturally only foresee and incentivize but cannot build. These structures—such as local organizations analyzing threats, consumer groups helping content creators, and communities of researchers—are the ones that give life to the DSA's tools. They need to be built from the bottom up by people, perhaps even locally in each member state. If their creation fails, the regulatory promises might turn out to be nothing more than glorious aspirations.

V. CONCLUSIONS

In 2023, content moderation continues to be a politically divisive topic in the United States. The Republican Party wants companies to moderate less content that is not prohibited by the legislature.[131] The Democratic Party wants them to moderate more of such content. The political currents have not yet swept Europe in a similar way, although the political situation is evolving.[132] While the two sides cannot agree on how to exercise content moderation discretion, they should be able to agree that legislative acts reinstating ex ante editors are in no one's interest.

The internet is a special medium that should not be regulated as broadcasting or newspapers. Content moderation discretion can only exist if providers have very limited liability for the distribution of the content of others. If liability is strict or close to strict, their discretion must morph into editorial discretion because no one can offer digital spaces or tools for expression without vetting information in advance.

Running our digital services—ranging from social media and marketplaces to search engines—on the infrastructure of editorial control is impossible. Thus, what policymakers should aim for is to increase providers' accountability while keeping their liability limited. Platforms need more *accountability, not liability*. Their design practices should be subject to regulation without immediately expanding the underlying content laws.

131. *See* S.B. 7072, 2021 Leg. (Fla.); H.B. 20, 2021 Leg., 87th Sess. (Tex.).

132. Among the E.U. countries, only Polish conservatives introduced a bill similar to the United States' Florida and Texas proposals. The Polish bill was meant to protect against "censorship" by prohibiting moderation of legal content. *Law To Protect Poles From Social Media "Censorship" Added To Government Agenda*, NOTES FROM POLAND (Oct. 5, 2021), https://notesfrompoland.com/2021/10/05/law-to-protect-poles-from-social-media-censorship-added-to-government-agenda/. However, the bill was never adopted. In the UK Online Safety Bill, the controversy around "lawful but harmful content" for adults led a new prime minister, Rishi Sunak, to drop the clause and only rely on empowerment obligation and extension of some offences.

Because non-editorial content lacks editors, some think it will also always lack *trust*. This leads policymakers to push for tighter content standards or even editorial discretion. However, there are ways to inject trust into the ecosystem without abandoning its decentralized character. The solution of the Trusted Content Creators,[133] for instance, draws entirely on the principles of *shared burden* and *ecosystem solutions*. Instead of banning or suppressing that what is not trusted, TCC rewards trusted content by asking providers to give extra benefits to those content creators who self-organize and commit to abide by their own shared norms. Decentralization is not the antithesis of trust.

Similarly, there are many ways to overcome different views on *how* to exercise content moderation discretion over legal content. The *user-empowerment* principle shows the way for a middle ground between two positions on how to exercise content moderation. It invites policymakers to think about solutions that delegate the choice of what legal content to display from advertisers or providers to individuals. The legislature can also facilitate user choice by making the underlying markets more competitive[134] or open up the content moderation experiences within dominant services to more alternatives.[135]

People voting with their feet show that they are interested in non-editorial content much more than they are in editorial content. Among the top fifty visited websites on the internet globally, the great majority rely on users—other people—to generate the content.[136] It seems that humans are primarily interested in what other humans have to say. No one can beat the educating and entertaining power of crowds. While we often fret about issues of the legality and trustworthiness of such content, only a few think the solution is to go back to the age of editorial media.

The proposed five principles offer common ground for liberal democracies to think about the challenges of our day without sacrificing what we have gained: an inclusive, decentralized and open global communication network.

133. Martin Husovec, *Trusted Content Creators*, LSE LAW POLICY BRIEFING PAPER NO. 52, (2022), https://ssrn.com/abstract=4290917.

134. This is the approach taken by the Digital Markets Act. *See* Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828, 2022 O.J. (L 265).

135. *See* Fukuyama et al., *supra* note 115 (proposing middle-ware).

136. *See List of Most-visited Websites*, WIKIPEDIA, https://en.wikipedia.org/wiki/List_of_most_visited_websites (last accessed Sept. 26, 2023).