Platform-Related Harms

Jordan Famularo

Introduction

Digital platforms have given rise to a lexicon of harms: digital harm, online harm, data harm, cyber harm, algorithmic harm, automated decision-making harm, and harmful content. These terms have proliferated in academic and civil society conversations about platforms with increasing momentum since 2017. Though the semantic richness of these terms might go unnoticed by the casual reader, their use in platform governance discussions provides a signal to policy and legal decision makers. These authorities should heed collective and societal harms, in addition to individual harms, in order to align with social norms. For purposes of this essay, I use the term "platform-related harms" to refer to the vocabulary of harms that scholars and advocates associate with platforms.

There are two fundamental reasons why the language of harm used by academia and civil society requires direct attention from policy makers and law makers. First, such language to denote platform-related harm is conceptually expansive and resists the longstanding legal definition of harm (i.e., a wrongful setback or thwarting of an interest).[1] Second, the same discourse reveals gaps between the law and societal expectations about acceptable platform behavior.

This essay clarifies discourse on platform-related harm by distilling a cross-cultural synopsis of vocabulary and rhetoric in sources published from 2014 to 2022 across four continents. To achieve this, it uses critical discourse analysis as framework,[2] integrating discursive patterns drawn

[1] *See* 1 JOEL FEINBERG, THE MORAL LIMITS OF CRIMINAL LAW: HARM TO OTHERS 34 (1987).

[2] *See* NORMAN FAIRCLOUGH, CRITICAL DISCOURSE ANALYSIS (1995); Norman Fairclough & R. Wodak, *Critical Discourse Analysis*, in 2 DISCOURSE AS SOCIAL INTERACTION 258 (Teun A. Van Dijk ed., 1997); Cynthia Hardy, *Researching Organizational Discourse*, 31 INT'L

from civil society, law, policy, social sciences, cybersecurity, and media studies. In doing so, it illuminates perspectives on identifying, characterizing, and managing relationships between platforms and harms. In turn, it reveals a new understanding of what platform-related harms are. They may be either tangible or intangible. They may be practices, outcomes, or media. And their effects happen at individual, collective, and systemic levels.

1. Expansive Concepts of Harm

What has harm become in the platform age? Harm has long been treated in sociology, psychology, and philosophy not as an objective condition but as a collectively negotiated concept. [3] It is shaped across time and place by social norms and local contexts. After the advent of the Internet, harm's connection to collective sense-making processes has afforded new meanings for platform-related harm. These meanings arise from expectations in communities[4] and may be at odds with legal concepts of harm. Three core features contribute to this wide conceptual ground.

First, platform-related harm is relative rather than absolute. This is because a behavior, outcome, or digital content might be perceived as harmful in one place but not the next, and to one person but not another. For instance, online hate speech might be harmful to targeted groups, but the verbal or visual frame of a specific post may affect whether an audience receives it as satirical or harmful, raising risks that content

---

STUD. MGMT. & ORG. 25 (2001); NORMAN FAIRCLOUGH, LANGUAGE AND POWER (3d ed. 2015). My discourse analysis in this essay is just one possible reading of a complex body of sources.

[3] *See, e.g.*, Herbert Blumer, *Social Problems as Collective Behavior*, 18 SOC. PROBS. 298 (1971); STEPHEN WILKINSON, BODIES FOR SALE: ETHICS AND EXPLOITATION IN THE HUMAN BODY TRADE (2003); BEYOND CRIMINOLOGY: TAKING HARM SERIOUSLY (P. Hillyard, C. Pantazis, S. Tombs & D. Gordon eds., 2004) (discussing a theory of harm irrespective of the Internet); Nick Haslam, *Concept Creep: Psychology's Expanding Concepts of Harm and Pathology*, 27 PSYCH. INQUIRY 1 (2016) (giving examples of studies that observe but do not theorize ways that harm differs across contexts); MARY KATE MCGOWAN, JUST WORDS: ON SPEECH AND HIDDEN HARM (2019); Nick Haslam et al., *Harm Inflation: Making Sense of Concept Creep*, 31 EURO. REV. SOC. PSYCH. 254 (2020).

[4] *E.g.*, Robert S. Tokunaga, *Following You Home From School: A Critical Review and Synthesis of Research on Cyberbullying Victimization*, 26 COMPUTS. HUMAN BEHAV. 277 (2010); Anastasia Powell & Nicola Henry, *Towards Equal Digital Citizenship*, *in* SEXUAL VIOLENCE IN A DIGITAL AGE 237, 242 (2017); Rahul Sinha-Roy & Matthew Ball, *Gay Dating Platforms, Crimes, and Harms in India: New Directions for Research and Theory*, 32 WOMEN & CRIM. JUSTICE 49 (2021).

moderation decisions will over-police or under-address harm.[5] Harm's relativity receives special emphasis in Indigenous[6] approaches to technology governance that treat harm as a symptom of imbalance in systems that shape society and the world. Analysis of sources centered on Indigenous perspectives for this essay found that they focus on twin aspects of harm—its systemic underpinnings and different appearances across occurrences.[7] For example, technology's disruption of respect between generations of people, or interference in the kinship network between machines and people, result in hierarchies that set up conditions for harm in local places and contexts.[8] Such claims therefore stress that social norms and localities are necessary to harm's meaning and relativity.

---

[5] *See* Libby Hemphill, *Very Fine People: What Social Media Platforms Miss About White Supremacist Speech*, ANTI-DEFAMATION LEAGUE 13 (May 3, 2022), https://www.adl.org/language-of-white-supremacy.

[6] There is no universally agreed upon terminology for referring to the many diverse groups who self-identify historical continuity with pre-colonial and/or pre-settler societies. This essay uses the term "Indigenous" to refer to all peoples and groups who identify as such, although they may use alternative designations, transliterated, for example, as Aboriginal or First Nations.

[7] *See Ownership, Control, Access, and Possession (OCAP™): The Path to First Nations Information Governance*, FIRST NATIONS INFO. GOVERNANCE CTR. (May 23, 2014), https://achh.ca/wp-content/uploads/2018/07/OCAP_FNIGC.pdf; FIRST NATIONS INFO. GOVERNANCE CTR., *Barriers and Levers for the Implementation of OCAP™*, 5 INT'L INDIGENOUS POL'Y J. 1 (2014); Inge Kral, *Shifting Perceptions, Shifting Identities: Communication Technologies and the Altered Social, Cultural and Linguistic Ecology in a Remote Indigenous Context*, 25 AUSTL. J. ANTHROPOLOGY 171 (2014); Petronella Vaarzon-Morel, *Pointing the Phone: Transforming Technologies and Social Relations among Warlpiri*, 25 AUSTL. J. ANTHROPOLOGY 239 (2014); Jason Edward Lewis, Noelani Arista, Archer Pechawis & Suzanne Kite, *Making Kin with the Machines*, 3.5 J. DESIGN & SCI. (2018), https://doi.org/10.21428/bfafd97b; Bronwyn Carlson & Ryan Frazer, *Social Media Mob: Being Indigenous Online*, MACQUARIE U. (Mar. 2018), https://research-management.mq.edu.au/ws/portalfiles/portal/85013179/MQU_SocialMediaMob_report_Carlson_Frazer.pdf; Bronwyn Carlson & Ryan Frazer, *Cyberbullying and Indigenous Australians: A Review of the Literature*, ABORIGINAL HEALTH & MED. RSCH. COUNCIL NEW SOUTH WALES (Sept. 2018), https://researchers.mq.edu.au/files/92634728/MQU_Cyberbullying_Report_Carlson_Frazer.pdf; *Indigenous Protocol and Artificial Intelligence*, INDIGENOUS PROTOCOL & ARTIFICIAL INTEL. WORKING GRP. (Jason Edward Lewis ed., Jan. 30, 2020), https://spectrum.library.concordia.ca/id/eprint/986506/7/Indigenous_Protocol_and_AI_2020.pdf; Ashley Cordes, *Meeting Place: Bringing Native Feminisms to Bear on Borders of Cyberspace*, 20 FEMINIST MEDIA STUD. 285 (2020); Stephanie Russo Carroll et al., *Operationalizing the CARE and FAIR Principles for Indigenous Data Futures*, 8 SCI. DATA 1 (2021); INDIGENOUS DATA SOVEREIGNTY AND POLICY (Maggie Walter et al. eds. 2021); Hēmi Whaanga & Paora Mato, *The Indigenous Data Footprint*, *in* ROUTLEDGE HANDBOOK OF CRITICAL INDIGENOUS STUDIES 447 (Brendan Hokowhitu et al. eds., 2020).

[8] *See* Lewis, Arista, Pechawis & Kite, *supra* note 7; *Indigenous Protocol and Artificial Intelligence*, *supra* note 7.

Second, notions of platform-related harm in academic and civil society sources cover wide conceptual ground by including consequences in multiple registers at the levels of individual, group, and society. Platform-related harm may have reinforcing effects beyond the sum of discrete setbacks to individual Internet users or single firms.[9] This is a fundamental reason to extend the concept of platform-related harm outside of traditional legal notions of a wrongful injury. For instance, a study of automated selection processes on social media platforms claims that "algorithmic harm" applies across persons, markets, and society as a whole.[10] The study exemplifies a common pattern in the discourse whereby scholars recognize harm's impacts on individuals, groups, and society.[11] Indeed, the buildup of micro-level harms into meso- and macro-level impacts is a notable theme in debates about platform governance,[12] which echo arguments that implicate the Internet in collective and social harms irrespective of blaming platforms specifically.[13] To mitigate or prevent certain aggregates of harms that platforms enable, policies fall short if they simply pinpoint single occurrences. This is because technical architectures and business logics give shape to exposure risks and cascading effects that are more synergistic than additive, according to some scholars.[14] Different registers of platform-related harm, scaled from individual to societal, are interconnected and difficult to sever from each other.

Third, academia and civil society recognize platform-related harm in a wide range of organizational practices, effects, human expression, and

---

[9] *See infra* notes 31-43 and accompanying text.

[10] *See* Florian Saurwein & Charlotte Spencer-Smith, *Automated Trouble: The Role of Algorithmic Selection in Harms on Social Media Platforms*, 9 MEDIA & COMMC'N 222, 223 (2021) ("[W]e use the term 'algorithmic harm' to describe harmful or negative effects upon individuals, markets, and society caused in part or in full by the use of algorithms.").

[11] *See* Powell & Henry, *supra* note 4; *see also infra* notes 33–34, 36–43, 49 and accompanying text.

[12] *See* Powell & Henry, *supra* note 4, at 253; Luke Price, *Platform Responsibility for Online Harms: Towards a Duty of Care for Online Hazards*, 13 J. MEDIA L. 238, 255 (2022).

[13] *E.g.*, Jane Bailey, *Confronting Collective Harm: Technology's Transformative Impact on Child Pornography*, 56 U. New Brunswick L.J. 65 (2007); Alessandro Manteloro, *From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era*, in GROUP PRIVACY: NEW CHALLENGES OF DATA TECHNOLOGIES 139 (Linnet Taylor, Luciano Floridi & Bart van der Sloot eds., 2017); Jay P. Kesan and Carol M. Hayes, *Liability for Data Injuries*, 2019 U. ILL. L. REV. 295 (2019); ANITA LAVORGNA, INFORMATION POLLUTION AS SOCIAL HARM: INVESTIGATING THE DIGITAL DRIFT OF MEDICAL MISINFORMATION IN A TIME OF CRISIS (2021).

[14] *See* Powell & Henry, *supra* note 4, at 253; THE EMERALD INTERNATIONAL HANDBOOK OF TECHNOLOGY-FACILITATED VIOLENCE AND ABUSE (Jane Bailey, Asher Flynn & Nicola Henry eds., 2021); Price, *supra* note 12, at 255.

media that do not map onto injury, as harm is traditionally known in law.[15] Policy-makers and others who wish to grasp the variety of phenomena that qualify as harm need to navigate an array of classification systems in the form of dictionaries, catalogs, and structured lists created by scholars and civil society practitioners. The discourse is rife with taxonomies, typologies, and registers that describe platform-related harms primarily by way of categories, which may be a sign that theory is seeking to keep up with a rapidly changing sociotechnical landscape. Some researchers propose organizing certain kinds of conduct, such as information-sharing practices or incendiary speech, into an overarching harm.[16] Others arrange specific effects, such as damage from an attack on digital infrastructure or an individual's loss of opportunity, into an overarching harm.[17] Still others distinguish harmful digital content by type such as hate, doxxing, and extremism.[18] classifications of harms included in the *Digital Harms Dictionary* categorizes harms into practice types such as information collection, information sharing, and computational modeling;[19] a "taxonomy" of organizational cyber harms that divides effects into economic, psychological, physical/digital, reputational, and social/societal;[20] and a "typology" of online content that can potentially pose risk of harm to users or be illegal in certain jurisdictions such as fake accounts, unauthorized intimate images, and disinformation.[21] Reviewing

[15] *See* FEINBERG, *supra* note 1.

[16] *E.g.*, *Digital Harms Dictionary 2.0*, INTERNET SAFETY LABS (June 8, 2021), https://internetsafetylabs.org/wp-content/uploads/2021/10/me2ba-digital-harms-dictionary-v2.0-iii.pdf; Tatjana Scheffler, Veronika Solopova, & Mihaela Popa-Wyatt, *The Telegram Chronicles of Online Harm*, 7 J. OPEN HUMANS. DATA (2021), at 1, https://storage.googleapis.com/jnl-up-j-johd-files/journals/1/articles/31/submission/proof/31-1-572-1-10-20210705.pdf.

[17] *E.g.*, *Unfairness by Algorithm: Distilling the Harms of Automated Decision-Making*, FUTURE PRIV. F. (Dec. 11, 2017), https://fpf.org/blog/unfairness-by-algorithm-distilling-the-harms-of-automated-decision-making; Joanna Redden, Jessica Brand & Vanesa Terzieva, *Data Harm Record (Updated)*, DATA JUST. LAB (Aug. 2020), https://datajusticelab.org/data-harm-record; Ioannis Agrafiotis, Jason R. C. Nurse, Michael Goldsmith, Sadie Creese, & David Upton, *A Taxonomy of Cyber-Harms: Defining the Impacts of Cyber-Attacks and Understanding How They Propagate*, 4 J. CYBERSEC. (2018), at 1, https://doi.org/10.1092/cybsec/tyy006; Saurwein & Spencer-Smith, *supra* note 10.

[18] *See Content & Jurisdiction Program: Operational Approaches: Norms, Criteria, Mechanisms*, INTERNET JURISDICTION & POL'Y NETWORK (Apr. 2019), https://www.internetjurisdiction.net/uploads/pdfs/Papers/Content-Jurisdiction-Program-Operational-Approaches.pdf; Michele Banko, Brendon MacKeen, & Laurie Ray, *A Unified Taxonomy of Harmful Content*, Proc. Fourth Workshop on Online Abuse & Harms 134-35 (Nov. 20, 2020). https://aclanthology.org/2020.alw-1.16.pdf.

[19] *Digital Harms Dictionary 2.0*, *supra* note 16.

[20] Agrafiotis, Nurse, Goldsmith, Creese, & Upton, *supra* note 17.

[21] *Content & Jurisdiction Program: Operational Approaches: Norms, Criteria, Mechanisms*, *supra* note 18.

these classification initiatives together, one begins to see that platform-related harm diverges from injury by occupying a spectrum that includes—among many phenomena—technologically enabled practice, effects which may be reputational or psychological, upsetting of social cohesion, and characteristics of online content.

2. What the Law Misses

The extralegal and norms-based features of platform-related harm have not gone unnoticed in policy and legal circles. Scholars have pointed out three prominent reasons why legal notions of harm fail to fully capture norms-based notions of harm. First, many areas of intermediary liability law across doctrines and jurisdictions have origins in tort law, which is most effective when harm is immediate and concrete, and it has so far failed to make actionable some phenomena perceived as harms and associated with platforms by civil society and academics.[22] For instance, victims of cyber mobs on social media platforms turn to the companies to shut down the abuse, but section 230 of the U.S. Communications Decency Act provides immunity from liability for user-generated content.[23]

Second, when latent harms are difficult to match causally with precise wrongs, tort law is inadequate for deterrence or compensation.[24] This problem crystallizes in Internet defamation cases against online intermediaries where, as Kylie Pappalardo and Nicolas Suzor explain for the Australian context, "the law on the distinction between 'active' 'publishing' and 'conduct that amounts only to the merely passive facilitation of disseminating defamatory matter' is still not well developed."[25]

---

[22] *See* Omri Ben-Shahar, *Data Pollution*, 11 J. LEGAL ANALYSIS 104, 124-27 (2019); Kylie Pappalardo & Nicolas Suzor, *The Liability of Australian Online Intermediaries*, 40 SYDNEY L. REV. 469 (2018); Jay P. Kesan & Carol M. Hayes, *Liability for Data Injuries*, 2019 U. ILL. L. REV. 295 (2019); Danielle Keats Citron & Daniel J. Solove, *Privacy Harms*, 102 B.U. L. REV. 793 (2022).

[23] *See* DANIEL J. SOLOVE, THE FUTURE OF REPUTATION: GOSSIP, RUMOR, AND PRIVACY ON THE INTERNET 152 (2007); Danielle Keats Citron, *Cyber Civil Rights*, 89 B.U. L. REV. 61, 116 (2009); DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE 170-71 (2014).

[24] *See* On the other hand, courts have occasionally recognized indirect causation of harm by platforms. In at least one U.S. federal case, *Perkins v. LinkedIn*, 53 F. Supp. 3d 1222 (N.D. Cal. 2014), the court held that the platform caused reputational harm to users indirectly by asking their email contacts to connect on the site without permission from individual members. For further discussion, *see* Citron and Solove, *supra* note 24, at 838-39.

[25] Pappalardo & Suzor, *supra* note 22.

Third, systems of procedural law have provided courts with means to deny standing to plaintiffs due to insufficient demonstration of harm.[26] In *In re Google, Inc. Privacy Policy Litigation*,[27] plaintiffs sued Google for using their personal data in ways different than what the company communicated, but the court found that they lacked standing because they failed to show how Google's "use of the information deprived the plaintiff of the information's economic value,"[28] as Daniel Solove and Danielle Keats Citron point out.[29]

These gaps in intermediary liability, tort, and procedural law suggest why it matters that policy makers and law makers recognize the distinction between legal and social norms-based concepts of harm. From the examples just given—abusive user-generated content, Internet defamation cases, and denial of standing because misuse of personal data fails to demonstrate harm sufficiently—society sees harm, but the law does not.

Some governance shortfalls arise from law and policy that inadequately address system-level problems, even if they provide protections for individuals. To underscore collective and societal harms that remain largely unchecked by current legal and policy regimes, scholars have developed a group of interrelated analogies. Until now, the legal and policy literatures have largely overlooked the existence and value of these analogies.

3. Exposing Supra-Individual Harms

A little-noticed linguistic theme in academic and civil society conversations about platform governance has recently become salient: environmental and ecological analogies. These describe platform-related harms' dynamics and effects. The benefits for rhetoric and analysis are

---

[26] *See* Daniel J. Solove & Danielle Keats Citron, *Risk and Anxiety: A Theory of Data Breach Harms*, 96 TEX. L. REV. 737 (2018); Ryan Calo, *A Long-Standing Debate: Reflections on* Risk and Anxiety: A Theory of Data Breach Harms *by Daniel Solove and Danielle Keats Citron*, 96 TEX. L. REV. ONLINE 59 (2018), https://digitalcommons.law.uw.edu/faculty-articles/410; Daniel J. Solove & Danielle Keats Citron, *Standing and Privacy Harms: A Critique of* TransUnion v. Ramirez, 101 B.U. L. REV. ONLINE 62 (2021), https://www.bu.edu/bulawreview/files/2021/07/SOLOVE-CITRON-2.pdf; Note that in the latter two articles, Solove and Citron argue that U.S. federal courts have interfered with states' provision of statutory private rights of action by way of standing doctrine.

[27] No. 12-cv-01382, 2013 WL 6248499 (N.D. Cal. Dec. 3, 2013).

[28] *Id.* at 5.

[29] Solove & Citron, *supra* note 26, at 850.

many, but the policy value has not fully unlocked, a point to which I will return.

Let's begin with the benefits. As platform governance discourse recognizes the Internet as a sociotechnical system that supports and shapes human activity, environmental and ecological analogies help scholars underscore the cumulative in a complex system. A multitude of living and non-living parts coexist, like a habitat or ecosystem. Where platforms have produced a *de facto* public environment, societal harms are a serious concern.[30] When platforms' data analytics demarcate groups and make decisions based on them, collective harm is a lens for bringing discriminatory data processing into focus.[31] A single instance of harm can be relatively benign, but if it proliferates cumulatively or systemically, then it can be perceived to impose harm at collective and/or societal levels. This point is insufficiently addressed, critics say, by legal and policy approaches that prioritize protections for individuals, or those that foreclose grounds to challenge the harm if individual harms are marginal or small.[32]

Observing that collective and societal harms are not always simply a sum of individual harms,[33] some scholars have highlighted in the context of platform governance a need to rethink dimensions of legal systems that prioritize individual rights and remedies.[34] Some proposals target harms of artificial intelligence (AI) systems used by platforms, which prompted one critic to observe that EU law currently hinges on private enforcement,

---

[30] See, for example, the European Union's Digital Services Act, recital 137, which states: "Given the importance of very large online platforms or very large online search engines, in view of their reach and impact, their failure to comply with the specific obligations applicable to them may affect a substantial number of recipients of the services across different Member States and may cause large societal harms, while such failures may also be particularly complex to identify and address." 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), 2022 O.J. (L 277) 1, 36 [hereinafter Digital Services Act].

[31] *See* GROUP PRIVACY: NEW CHALLENGES OF DATA TECHNOLOGIES (Linnet Taylor, Luciano Floridi & Bart van der Sloot eds., 2017).

[32] *See id*.; David J. Bier, *Integrating Integrity: Confronting Data Harms in the Administrative Age*, 99 B.U. L. REV. 1799 (2019); Lauren E. Elrick, *The Ecosystem Concept: A Holistic Approach to Privacy Protection*, 35 INT'L REV. L., COMPUTS. & TECH. 24 (2021); Bart van der Sloot & Sascha van Schendel, *Procedural Law for the Data-Driven Society*, 30 INFO. & COMMC'NS TECH. L. 304 (2021); Citron & Solove, *supra* note 26.

[33] *See* CHRISTOPHER KUTZ, COMPLICITY: ETHICS AND LAW FOR A COLLECTIVE AGE (1st ed. 2000).

[34] *See* Natalie A. Smuha, *Beyond the Individual: Governing AI's Societal Harm*, 10 INTERNET POL'Y REV. (Sept. 30, 2021), at 1, https://policyreview.info/articles/analysis/beyond-individual-governing-ais-societal-harm; Powell & Henry, *supra* note 4; Price, *supra* note 12.

relying on individuals to contest potential harms and failing to address some societal harms. The author presses the point by drawing an analogy between societal-level algorithmic harm and environmental harm.[35]

Examples of cumulative platform-enabled harms that snowball into collective or societal harms, according to scholars and advocates, include manipulation that exploits human bias or vulnerability,[36] technology-facilitated sexual violence,[37] hateful content,[38] and harmful algorithmic gatekeeping.[39] For example, among the criticisms of the proposed EU Artificial Intelligence Act[40] is its failure to address social harms such as "harm to . . . democratic societies" that could arise from AI systems that were, from an *ex-ante* perspective, not classified as high-risk under the current proposal but turn out to have severe and detrimental impacts on societies.[41] In the context of debate about governing platforms through AI regulation, supra-individual harm motivates analogies with environmental harm.[42]

Legal and policy scholars have underscored that platforms impose externalities, familiar from environmental discourse, on individuals and groups. Like polluters that externalize costs on ecosystems and their inhabitants, platforms externalize costs on communities and people, according to privacy and data governance experts.[43] Proposed mitigations for platform-related societal harms target the logic of externalities by

---

[35] Smuha, *supra* note 34, at 4 ("[H]ow can we reconcile the need to protect societal interests adversely impacted by AI in the context of a legal system that primarily focuses on individual rights and remedies? . . . An analogy can . . . be drawn with environmental harm, which likewise encompasses a societal dimension that cannot always be reduced to demonstrable individual harm.").

[36] *See* Solove and Citron, *supra* note 26, at 847.

[37] *See* Powell & Henry, *supra* note 4, at 253.

[38] *See* Hemphill, *supra* note 5, at 14.

[39] *See* Zeynep Tufekci, *Algorithmic Harms Beyond Facebook and Google: Emergent Challenges of Computational Agency*, 13 COLO. TECH. L.J. 203 (2015).

[40] *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206 final Apr. 21, 2021).

[41] *Draft AI Act: EU Needs to Live Up to its Own Ambitions in Terms of Governance and Enforcement*, ALGORITHM WATCH 4 (Aug. 2021), https://algorithmwatch.org/en/wp-content/uploads/2021/08/EU-AI-Act-Consultation-Submission-by-AlgorithmWatch-August-2021.pdf.

[42] *See* Smuha, *supra* note 34.

[43] *See* James P. Nehf, *Recognizing the Societal Value in Information Privacy*, 78 WASH. L. REV. 1 (2003); Dennis D. Hirsch & Jonathan H. King, *Big Data Sustainability: An Environmental Management Systems Analogy*, 72 WASH. & LEE L. REV. ONLINE 406 (2016), https://scholarlycommons.law.wlu.edu/cgi/viewcontent.cgi?article=1039&context=wlulr-online; Salomé Viljoen, *A Relational Theory of Data Governance*, 131 YALE L.J. 573 (2021).

adapting solutions from EU and American environmental law.[44] For instance, a public enforcement model similar to environmental law's approaches to pollution could redress unwanted data transmissions such as leaks and inappropriate sharing.[45]

Furthermore, the concept of externality can clarify bystander problems in platform-related harm, according to some privacy scholars.[46] Platforms' dependence on data-driven enterprise is the focus because of its potential effects on third parties who do not directly participate in its data transactions but are still exposed to the broader data brokerage environment. Third parties can experience loss of opportunity, differential access, or discrimination when the platform's data-sharing activities enable inferences about them based on traits, social contacts, or histories that they share with people who are the platform's data subjects.[47] Such losses and impairments are commonly known in platform governance discourse as a subset of algorithmic harms or automated decision-making harms, which can affect collectives and individuals.[48] Some are prohibited by law, whereas others are regarded as simply unfair, giving rise to debate about how reform should develop to cover the latter.[49]

The above examples of analogies with ecology and the environment show that fundamental policy assumptions need reevaluation. Scholars deploy such analogies to assess the viability of premises that are grounded in Western ideas of personhood and individualism, and to point out that collective and social harms have so far proven largely intractable.

The practical value of these analogies for platform governance is little realized, but still possible. This circumstance may be conditional: a point

---

[44] *See* Ben-Shahar, *supra* note 22; Lina M. Khan & David E. Pozen, *A Skeptical View of Information Fiduciaries*, 133 HARV. L. REV. 497 (2019); Smuha, *supra* note 34.

[45] *See* Ben-Shahar, *supra* note 22, at 131-48.

[46] *See* Dennis D. Hirsch, *Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law*, 41 GA. L. REV. 1 (2006); A. Michael Froomkin, *Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements*, 2015 U. ILL. L. REV. 1713 (2015).

[47] *See* Ben-Shahar, *supra* note 22, at 115-16; Dirk Bergemann, Alessandro Bonatti & Tan Gan, *The Economics of Social Data*, ARXIV (Nov. 20, 2021), at 2ff, https://arxiv.org/pdf/2004.03107v1.pdf.

[48] *E.g.*, FUTURE OF PRIVACY FORUM, *supra* note 17; Saurwein & Spencer-Smith, *supra* note 10, at 227; Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2019 COLUM. Bus. L. REV. 494 (2019).

[49] *E.g.*, *Unfairness by Algorithm: Distilling the Harms of Automated Decision-Making*, *supra* note 17; Rebecca Kelly Slaughter, Janice Kopec, & Mohamad Batal, *Algorithms and Economic Justice: A Taxonomy of Harms and a Path Forward for the Federal Trade Commission*, 23 YALE J.L. & TECH. SPECIAL ISSUE 1 (2021).

in the unfolding history of platform-related harms rather than reflective of these analogies' real capacity for catalyzing policy change. In the current moment, the gap between prevailing individual rights-based regimes and observable collective and societal harms remains a significant challenge in platform governance. In short, there is a mismatch between discourse and policy. Few policy innovations foreground collective and societal harms, despite scholarly and civil society conversations that prioritize them. One exceptional area of progress, however, is collective data privacy in Canada, where Indigenous communities that have adopted the First Nations Information Governance Centre OCAP® principles[50] have passed their own privacy laws. As Tahu Kukutai and Donna Cormack explain, Indigenous privacy interests are intertwined with a normative and social system that emphasizes totality and interconnectedness, which diverges from settler (Western) notions of property, ownership, and privacy.[51] In a Western context, scholars have suggested that institutional hurdles impede a workable concept of group privacy as a countermeasure to harm: collectives might not be aware that they are at risk of platform-related harm or have legal instruments to seek remedies, while institutions in scope are insufficiently empowered to act and regulate.[52]

4. A Culture of Risk

Environmental and ecological language in platform governance discourse feeds into a certain orientation to risk-based regulation. Risk-based regulation has come into favor in environmental regulation and now regulation of digital society and digital markets.[53] Sociologists Ulrich Beck and Anthony Giddens theorized that contemporary society is a "risk society," a term concerned with the transition from industrial society to

[50] OCAP® is a registered trademark of the First Nations Information Governance Centre (FNIGC). *See The First Nations Principles of OCAP®*, FIRST NATIONS INFO. GOVERNANCE CTR., https://fnigc.ca/ocap-training..

[51] Tahu Kukutai & Donna Cormack, *"Pushing the Space": Data Sovereignty and Self-Determination in Aotearoa NZ*, *in* INDIGENOUS DATA SOVEREIGNTY AND POLICY 29 (Maggie Walter et al. eds., 2021) (citing James Williams, Megan Vis-Dubar & Jens Weber, *First Nations Privacy and Modern Health Care Delivery*, 10 INDIGENOUS L.J. 101 (2011)).

[52] *See* Linnet Taylor, Bart van der Sloot & Luciano Floridi, *Conclusion: What Do We Know about Group Privacy?*, *in* GROUP PRIVACY: NEW CHALLENGES OF DATA TECHNOLOGIES, *supra* note 31, at 233.

[53] *See* Zohar Efroni, *The Digital Services Act: Risk-based Regulation of Online Platforms*, INTERNET POL'Y REV. (Nov. 16, 2021), https://policyreview.info/articles/news/digital-services-act-risk-based-regulation-online-platforms/1606.

the current era shaped much more by technological hazards.[54] The risk society is distinguished not only by distribution of " goods" (wealth) but more so by distribution of "bads" (technological hazards produced by society such as pollution, contamination, cyberattacks, and election interference). The risk society is also increasingly preoccupied with the future, which generates the notion of risk. In the contemporary period the principle of risk captures growing attention in regulation of digital society, exemplified for instance by the EU Digital Services Act, as Zohar Efroni observes.[55] The Digital Services Act imposes an obligation on social media platforms to perform risk assessments to uncover threats presented by illegal content and the effects on fundamental rights, civic discourse, elections, public security, and public health, among other topics.[56] The act is also marked by risk regulation mechanisms such as risk management, risk mitigation, audit, and reporting.[57]

Environmental and ecological language is useful for platform governance policy, but the full value will partly turn on the viability of risk as a decision-making tool for mitigating harm at group and population levels. Whereas environmental policy is serviced by the field of quantitative environmental risk analysis,[58] there is not yet a mature science of risk at this point in the concern about platform-related harms.[59] This poses an obstacle for those who wish to push environmental and ecological language toward policy interventions for platform-related harm. Platform governance needs to catch up with the fields of public health and ecology in regard to developing a scientific basis for approaching risk to human and societal well-being. Regulators, regulated entities, and social scientists who examine interrelations between platforms and society need to develop a more systematic understanding of novel risks.

In platform governance there is a discourse of risk but few methods for quantifying likelihood and severity, although research on the topic is beginning to emerge.[60] Exposure to risk of harm arising from platforms

[54] ULRICH BECK, RISK SOCIETY: TOWARDS A NEW MODERNITY (Mark Ritter, trans., 1992); Anthony Giddens, *Risk and Responsibility*, 62 MOD. L. REV. 1 (1999).

[55] Efroni, *supra* note 53.

[56] Digital Services Act, *supra* note 32.

[57] *See* Efroni, *supra* note 53.

[58] *See* Robert A. Field, Norman A. Eisenberg, & Keith L. Compton, QUANTITATIVE ENVIRONMENTAL RISK ANALYSIS FOR HUMAN HEALTH (2007).

[59] *See* Efroni, *supra* note 53.

[60] *See* Johanne Kübler et al., *The 2021 German Federal Election on Social Media: An Analysis of Systemic Electoral Risks Created by Twitter and Facebook Based on the Proposed EU Digital Services Act*, SUSTAINABLE COMPUTING LAB & VIENNA U. ECON. & BUS. (Aug.

plays a role in the governance debate,[61] so there is a culture of risk, but there are no widely accepted risk metrics. To comprehend the culture of risk, consider that in publicly funded research there have emerged risk classifications. To address children's risk of online harm, for example, an EU-funded project presents a four-part taxonomy composed of content risk, contact risk, conduct risk, and contract risk.[62] The proposed Online Safety Bill in the United Kingdom would require social media platforms in scope to publish risk assessments.[63] These examples suggest that risk identification and assessment are taking root in platform governance. Looking to the history of environmental policy for an analogue,[64] it is reasonable to say that more should be done to improve risk analysis for platform-related harms as a way to enhance understanding between policymakers and social scientists.

Risk in the discourse on platform-related harm, and the environmental and ecological language that gives it rhetorical backing, may gain more traction if decision makers incentivize the right quantitative research. Grants should go toward establishing a body of research to support statistically rigorous mapping between risks and realized harms. The goal should be to develop validated means to assess likelihood and severity, and to underscore with consequence modeling how these relate to collective and societal harms. What is the probability that distorted election news weakens social cohesion, for instance, and how severe is the cost? Platforms complicate questions such as these. Answers may

---

2021), https://www.sustainablecomputing.eu/wp-content/uploads/2021/10/DE_Elections_Report_Final_17.pdf.

[61] *E.g.*, Jamie-Lee Mooney, *Protecting Children from the Risk of Harm? A Critical Review of the Law's Response(s) to Online Child Sexual Grooming in England and Wales*, *in* MINDING MINORS WANDERING THE WEB: REGULATING ONLINE CHILD SAFETY 283 (S. van der Hof, B. van den Berg, & B.Schermer eds., 2014); *Towards an Internet Safety Strategy*, 5RIGHTS FOUND. 4-5 (2019), https://5rightsfoundation.com/uploads/final-5rights-foundation-towards-an-internet-safety-strategy-january-2019.pdf; Karoline Andrea Ihlebæk & Vilde Schanke Sundet, *Global Platforms and Asymmetrical Power: Industry Dynamics and Opportunities for Policy Change*, NEW MEDIA & SOC'Y (forthcoming) (manuscript at 1, 9), https://journals.sagepub.com/doi/10.1177/14614448211029662; Citron & Solove, *supra* note 24.

[62] The European Union's Horizon 2020 funding program, for example. supported work on online risks to children. *See* Sonia Livingstone & Mariya Stoilova, *The 4Cs: Classifying Online Risk to Children, Children Online: Research and Evidence*, (CO:RE) Short Report Series on Key Topics, LEIBNIZ-INSTITUT FÜR MEDIENFORSCHUNG & HANS-BREDOW-INSTITUT (2021), https://doi.org/10.21241/ssoar.71817.

[63] Draft Online Safety Bill 2021, CP 405, Dep't for Sci., Innovation & Tech. & Dep't for Culture, Media & Sport (May 21, 2021), https://www.gov.uk/government/publications/draft-online-safety-bill.

[64] *See* JOHN M. STONEHOUSE AND JOHN D. MUMFORD, SCIENCE, RISK ANALYSIS AND ENVIRONMENTAL POLICY DECISIONS (U.N. Env't Programme 1995).

require more funding and advocacy for studies that address quantification of risk.

Conclusion

Language about perceptions of platform-related harm offers a window into cultural paradigms that underlie seeing, assuming, asserting, and disputing links between platforms and harms. It thus provides a basis for understanding how and where there are mismatches between societal expectations and governance, based on civil society and academic voices examined here. Advocates and scholars have uncovered weakness in platform governance by exposing supra-individual harms with focused vocabulary, drawing comparisons with ecology and the environment.

This essay's main contribution is to suggest that decision-makers need to take steady steps toward policy and legal models that consider collective and societal harms, in addition to individual harms, if they wish to mirror evolving norms. It also sharpens the view on why purposeful change in the way that we talk about harms can transform some governance gaps from seemingly unworkable to actionable. Calling out the differences between the prevailing legal sense of harm and social norms around harm can shift decision-makers' energy toward governance that addresses these shortfalls.