Washington University Law Review

Volume 96
Issue 6 *Trust and Privacy in the Digital Age*

2019

Safe Social Spaces

Ari Ezra Waldman
Princeton University, Center for Information Technology Policy; New York Law School

SAFE SOCIAL SPACES

ARI EZRA WALDMAN*

ABSTRACT

Technologies that mediate social interaction can put our privacy and our safety at risk. Harassment, intimate partner violence and surveillance, data insecurity, and revenge porn are just a few of the harms that bedevil technosocial spaces and their users, particularly users from marginalized communities. This Article seeks to identify the building blocks of safe social spaces, or environments in which individuals can share personal information at low risk of privacy threats. Relying on analogies to offline social spaces—Alcoholics Anonymous meetings, teams of coworkers, and attorney-client relationships—this Article argues that if a social space is defined as an environment characterized by disclosure, then a safe social space is one in which disclosure norms are counterbalanced by equally as powerful norms of trust that are both endogenously designed in and backed exogenously by law. Case studies of online social networks and social robots are used to show how both the design and law governing technosocial spaces today not only do not support trust, but actively undermine user safety by eroding trust and limiting the law's regulatory power. The Article concludes with both design and law reform proposals to better build and protect trust and safe social spaces.

* Microsoft Visiting Professor of Information Technology, Princeton University, Center for Information Technology Policy. Professor of Law and Founding Director, Innovation Center for Law and Technology, New York Law School. Affiliate Fellow, Information Society Project, Yale Law School. Ph.D., Columbia University; J.D., Harvard Law School. Thank you to Danielle Keats Citron, Elisa D'Amico, Mary Anne Franks, Bradley Greenberg, Woodrow Hartzog, Margaret Hu, Leslie John, Kate Klonick, Amanda Levendowski, Doug Lichtman, Neil Richards, Andrew Santa Ana, and Paul Schwartz. Maverick James provided essential research assistance. A version of this article was honored as the Deirdre G. Martin Memorial Lecture in Privacy at the University of Ottawa, Faculty of Law. And it is dedicated in the memory of Ian Kerr. It benefited greatly from comments and feedback from participants at the Internet Law Works in Progress Conference and from workshops and discussions with many colleagues. Small portions of this article are adapted from my book, *Privacy As Trust*, but my work here represents additional and refined thoughts based on new research. Special thank you to and to Neil Richards, Woody Hartzog, and Danielle Keats Citron. All errors are my own. I trust you will forgive me for them.

TABLE OF CONTENTS

INTRODUCTION

Our social interactions are mediated by technology. We chat with friends, read the news, and buy things on technosocial[1] platforms run by the likes of Facebook, Google, and Amazon. Alongside all of the benefits that kind of technology offers, it can also put our privacy and safety at risk. Scholars and media commentators have documented the rampant invasions

1. *See* Katherine J. Strandburg, *Home, Home on the Web and Other Fourth Amendment Implications of Technosocial Change*, 70 MD. L. REV. 614, 619 (2011) (coining the term "technosocial" to refer to the "intertwined effects of technological and social change"). This article uses "technosocial" to refer to technologies that are either themselves social or that foster social interaction.

of privacy,[2] gender-based harassment,[3] racism,[4] cyberstalking,[5] nonconsensual pornography,[6] and intimate surveillance[7] on digital social platforms. Prominent women and members of other marginalized groups are leaving these spaces.[8] That is not only regrettable; it is dangerous for democracy.[9] Even as scholars start to pay more attention to platform content moderation policies that ostensibly try to create safe and welcoming environments online,[10] things are not much better on the ground. This raises the question at the heart of this article: How can we make online social spaces safer?

Social spaces, as I am using the phrase, are multi-actor information-sharing environments.[11] They can be physical (chatting with a friend at a coffee shop or running into an acquaintance and her dog at the corner of First and Main), digital (texting with someone on an online social network),

2. *See, e.g.*, Sam Wolfson, *Amazon's Alexa Recorded Private Conversation and Sent it to Random Contact*, GUARDIAN (May 24, 2018, 6:09 PM), https://www.theguardian.com/technology/2018/may/24/amazon-alexa-recorded-conversation [https://perma.cc/NM8F-HUKB]; Issie Lapowsky, *Facebook Exposed 87 Million Users to Cambridge Analytica*, WIRED (Apr. 4, 2018, 5:43 PM), https://www.wired.com/story/facebook-exposed-87-million-users-to-cambridge-analytica/ [https://perma.cc/K3VS-6BYB].

3. *See* DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE (2014) (documenting a variety of forms of gender-based harassment and arguing for a civil rights agenda to combat them); Danielle Keats Citron, *Cyber Civil Rights*, 89 B.U. L. REV. 61 (2009) [hereinafter Citron, *Cyber Civil Rights*] (arguing that gender-based harassment online is a civil rights violation).

4. *See, e.g.*, JESSE DANIELS, CYBER RACISM (2009); Brandon A. Robinson, *"Personal Preference" as the New Racism: Gay Desire and Racial Cleansing in Cyberspace*, 2 SOC. RACE & ETHNICITY 317 (2015).

5. *See, e.g.*, Danielle Keats Citron, *Spying Inc*., 72 WASH. & LEE L. REV. 1243, 1248–50 (2015) (exploring the federal and state criminal laws that punish and deter businesses trafficking in devices that are primarily useful for surreptitious interception of electronic communications).

6. *See* Ari Ezra Waldman, *Law, Privacy, and Online Dating: "Revenge Porn" in Gay Online Communities*, 44 L. & SOC. INQUIRY, 2019, https://www.cambridge.org/core/services/aop-cambridge-core/content/view/BCCE05CF25AA4C2E05CCF8D64980E839/S0897654618000291a.pdf/law_privacy_and_online_dating_revenge_porn_in_gay_online_communities.pdf [https://perma.cc/L3HB-BB7T] [hereinafter Waldman, *Law, Privacy, and Online Dating*] (documenting the phenomenon of revenge porn in gay male online communities).

7. *See, e.g.*, Diana Freed et al., *Digital Technologies and Intimate Partner Violence: A Qualitative Analysis with Multiple Stakeholders*, ACM ON HUMAN-COMPUTER INTERACTION, 2017, http://www.nix dell.com/papers/digital-technologies-intimate.pdf [https://perma.cc/XJ6U-K5VD].

8. *See* Catherine Piner, *Feminist Writer Jessica Valenti Takes a Break from Social Media After Threat Against Her Daughter*, SLATE (July 28, 2016, 5:01 PM), https://slate.com/human-interest/2016/07/feminist-writer-jessica-valenti-takes-a-break-from-social-media-after-threat-against-her-daughter.html [https://perma.cc/246F-A72E].

9. *See* Danielle Keats Citron, *Law's Expressive Value in Combatting Gender Harassment*, 108 how tMICH. L. REV. 373, 391 (2009) [hereinafter Citron, *Law's Expressive Value*] (discussing some of the broader harms of gender-based harassment and the silencing of women that it causes, including entrenching traditional hierarchies and eroding the ability of women to contribute to society, generally).

10. *See* Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018) (chronicling the development of content moderation policies at Facebook and arguing that they reflect First Amendment norms).

11. *See* ERVING GOFFMAN, THE PRESENTATION OF SELF IN EVERYDAY LIFE 1–4 (1959) [hereinafter GOFFMAN, EVERYDAY LIFE].

or telephonic;[12] they can be big (a party or a megachurch) or small (a one-on-one meeting); they can involve the exchange of words (during a conversation over dinner) or body language (at a dance party or across the room at a tiresome meeting).[13] Spaces become *social* when they are constructed by persons engaged in information exchange.

As such, social spaces require us to navigate our privacy. Granted, privacy and sharing are creatures of context,[14] and different social contexts function on different disclosures.[15] But all social spaces operate with disclosure norms; that is, we all must share something.[16] Because sharing information involves some risk—disclosure inherently makes one vulnerable to others—social spaces require risk minimization mechanisms if they are to survive. Otherwise, we could not continue to share secrets with our best friends, confide in loved ones, engage in commerce, or express ourselves freely.[17] We would lose our sexual privacy,[18] our opportunities for

12. Although this understanding of social spaces is indebted to Goffman's work on the interaction among persons in public places, Goffman's research was focused exclusively on individuals in the "presence of others." *Id.* at 1. Goffman even defines "interaction" to mean "face-to-face interaction," as if there could be no other kind. *Id.* at 15. There is now a long literature applying Goffman's concept of impression management to online social life. *See, e.g.*, Liam Bullingham & Ana C. Vasconcelos, '*The Presentation of Self in the Online World': Goffman and the Study of Online Identities*, 39 J. INFO. SCI. 101 (2013) (similar to offline, face-to-face interactions, internet users re-create their offline self online, but engage in image management and persona editing); Julie E. Cohen, *Privacy, Visibility, Transparency, and Exposure*, 75 U. CHI. L. REV. 181, 197–98 (2008) (selective exposure, critical to Goffman's presentation of self, animates the different levels of privacy that some social networking sites provide); Jennifer L. Gibbs, Nicole B. Ellison & Rebecca D. Heino, *Self-Presentation in Online Personals: The Role of Anticipated Future Interaction, Self-Disclosure, and Perceived Success in Internet Dating*, 33 COMM. RES. 152 (2006) (examining conditional self-disclosure in the online dating context); Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1427 (2000) (arguing that information privacy allows us to construct Goffman's social facades).

13. You can communicate quite a bit of information through body language. *See, e.g.*, Lorenza Mondada, *Challenges of Multimodality: Language and the Body in Social Interaction*, 20 J. SOCIOLINGUISTICS 336, 340–41 (2016) (discussing the challenges of studying social interaction when language is only one modality); *see also* NONVERBAL BEHAVIOR AND COMMUNICATION (Aron W. Siegman & Stanley Feldstein eds., 2d ed. 2009).

14. *See, e.g.*, HELEN NISSENBAUM, PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE 137–38 (2010) (arguing that norms of appropriateness and information flow govern our expectations of privacy in different contexts).

15. *See* Georg Simmel, *The Sociology of Secrecy and of Secret Societies*, 11 AM. J. SOC. 441, 442–45, 463 (1906) (arguing that our relationships with different people differ because we share certain information with some and not with others); GOFFMAN, EVERYDAY LIFE, *supra* note 11, at 107, 112 (noting that forms of social interaction occur in the contexts appropriate for them, and what would be appropriate in one context might not be appropriate in another).

16. *See* GOFFMAN, EVERYDAY LIFE, *supra* note 11, at 64; *see also* GEORG SIMMEL, THE SOCIOLOGY OF GEORG SIMMEL 315–16, 326–29 (Kurt H. Wolff ed. & trans., 1950); SISSELA BOK, SECRETS: ON THE ETHICS OF CONCEALMENT AND REVELATION 20, 23 (1982).

17. *See* Neil Richards & Woodrow Hartzog, *Taking Trust Seriously in Privacy Law*, 19 STAN. TECH. L. REV. 431, 452–56 (2016).

18. *See* Danielle Keats Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1874 (2019) [hereinafter Citron, *Sexual Privacy*] (sexual privacy involves "the social norms (behaviors, expectations, and expectations) that govern access to, and information about, individuals' intimate lives.").

solitude,[19] and our freedom to develop and affirm our identities as we see fit.[20] Creating environments where these freedoms exist is, I argue, the role of trust, design,[21] and the law.[22] If social spaces are defined by information exchange, *safe* social spaces are environments of information exchange in which disclosure norms are counterbalanced by norms of trust backed endogenously by design and exogenously by law.

To suggest that the buildings blocks of safe social spaces are trust, design, and law, this article offers analogies.[23] Part I explores trust and its effects on social behavior. Part II then shows how three paradigmatic safe social spaces—Alcoholics Anonymous (AA) meetings, corporate teams, and attorney-client relationships—are all endogenously designed to foster

19. Solitude is an important value long coveted by privacy scholars. *See, e.g.*, Julie E. Cohen, *DRM and Privacy*, 18 BERKELEY TECH. L.J. 575, 576–79 (2003) [hereinafter Cohen, *DRM and Privacy*] (identifying intellectual privacy as an important value and noting its connection to privacy and solitude); Neil Richards, *Intellectual Privacy*, 87 TEX. L. REV. 387 (2008) (arguing that the ability to test out inchoate or unpopular ideas requires freedom from social surveillance); Samuel Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 196 (1890) ("The intensity and complexity of life, attendant upon advancing civilization, have rendered necessary some retreat from the world, and man, under the refining influence of culture, has become more sensitive to publicity, so that solitude and privacy have become more essential to the individual.").

20. *See* IRWIN ALTMAN, THE ENVIRONMENT AND SOCIAL BEHAVIOR: PRIVACY, PERSONAL SPACE, TERRITORY, CROWDING 49–50 (1975) ("when the permeability of these boundaries [to the self] is under the control of a person a sense of individuality develops."); *see also* Edward J. Bloustein, *Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser*, 39 N.Y.U. L. REV. 962, 973–74 (1964) (arguing that one who is subject to privacy intrusions is "less of a man, [and] has less human dignity").

21. The word "design" can mean many different things, from intentions (something is done "by design") to aesthetics (a room can be designed to be visually appealing). But for the purposes of this Article, I follow a broad definition from Don Norman, who wrote about design as a combination of affordances, constraints, and guideposts that direct behavior in useful ways. *See* DON NORMAN, THE DESIGN OF EVERYDAY THINGS 2, 9, 12–13 (1988). Sometimes, these are obvious: a wall in front of you redirects your path. Sometimes, these are subtler: spokes on street-level window sills discourage loitering. With respect to online spaces, I follow Woodrow Hartzog in his book, *Privacy's Blueprint*, which defines design as the "processes that create consumer technologies and the results of their creative processes instantiated in hardware and software." WOODROW HARTZOG, PRIVACY'S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF NEW TECHNOLOGIES 11 (2018) [hereinafter HARTZOG, PRIVACY'S BLUEPRINT].

22. *See* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE (1999) [hereinafter LESSIG, CODE] (noting that law, norms, markets, and technological architecture all govern conduct online). Trust is a norm. *See* Francis Fukuyama, *Differing Disciplinary Perspectives on the Origins of Trust*, 81 B.U. L. REV. 479, 480–81 (2001) [hereinafter Fukuyama, *Differing Disciplinary Perspectives*]. Design is architecture. *See* HARTZOG, PRIVACY'S BLUEPRINT, *supra* note 21, at 11.

23. *See, e.g.*, Larry Alexander, *Premises and Conclusions: Symbolic Logic for Legal Analysis: The Banality of Legal Reasoning*, 71 NOTRE DAME L. REV. 517 (1998); Scott Brewer, *Exemplary Reasoning: Semantics, Pragmatics and the Rational Force of Legal Argument by Analogy*, 109 HARV. L. REV. 923, 937 (1996); Ronald Dworkin, *In Praise of Theory*, 29 ARIZ. ST. L.J. 353 (1996); James R. Murray, *The Role of Analogy in Legal Reasoning*, 29 U.C.L.A. L. REV. 833 (1982); Frederic Schauer, *Precedent*, 39 STAN. L. REV. 571 (1987); Cass R. Sunstein, *On Analogical Reasoning*, 106 HARV. L. REV. 741 (1993). There are, of course, other relationships that offer informative analogies, including, in particular, the social relationship established between investigative reporters and their sources. The three discussed in this article are emblematic of safe social spaces.

the kind of trust, confidentiality, and discretion needed to facilitate disclosure.[24] And because society benefits from disclosures in each of these contexts,[25] the law exogenously supports designed-in norms of trust to ensure those spaces are safe for sharing personal, secret, or stigmatizing information.

Technosocial spaces, however, lack both endogenous and exogenous structures that support trust. Far from it. As I discuss in Part III, these spaces are actually designed to manipulate us and lull us into false senses of familiarity and confidence, thereby enticing risky disclosure. And they do so in a legal and regulatory void that leaves users unprotected and vulnerable to invasions of privacy and online harassment.

But it doesn't have to be that way. Technosocial spaces can learn from safe social spaces offline and reorient design and law to foster trust. Robust approaches to privacy- and safety-by-design can protect users from the inside,[26] and stronger legal responses to manipulative design and online harassment can restore trust when something goes wrong. These proposals are outlined in Part IV.

---

24. *See* ARI EZRA WALDMAN, PRIVACY AS TRUST: INFORMATION PRIVACY FOR AN INFORMATION AGE 54–60 (2018) [hereinafter WALDMAN, PRIVACY AS TRUST] (showing that trust among social actors is essential for sharing personal information).

25. For some examples of the scholarly literature on the benefits of sharing and disclosure in addiction recovery, please see, e.g., Irwin Altman, *Reciprocity of Interpersonal Exchange*, 3 J. THEORY SOC. BEHAVIOUR 249 (1973); Kathryn P. Davison, James W. Pennebaker & Sally W. Dickerson, *Who Talks? The Social Psychologist of Illness Support Groups*, 55 AM. PSYCHOLOGIST 205 (2000); Dalmas A. Taylor, *The Development of Interpersonal Relationships: Social Penetration Processes*, 75 J. SOC. PSYCH. 79 (1968). For the value generated by free sharing among corporate teams, please see, e.g., Amy C. Edmondson, *The Local and Variegated Nature of Learning in Organizations: A Group-Level Perspective, in* SOCIOLOGY OF ORGANIZATIONS: STRUCTURES AND RELATIONSHIPS 631 (Mary Goodwyn & Jody Hoffer Gittel eds., 2012) [hereinafter Edmondson, *Learning in Organizations*] (discussing how teams of coworkers work together and share information to enhance organizational learning); Morten T. Hansen, *The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge Across Organization Subunits*, 44 ADMIN. SCI. Q. 82, 105 (1999) (showing that complex information is difficult to transmit between corporate department teams); David Lazer & Allan Friedman, *The Network Structure of Exploration and Exploitation*, 52 ADMIN. SCI. Q. 667 (2007) (considering how network structure affects communication among teams and, ultimately, system performance). And to understand the benefits to society from the free flow of information between clients and attorneys, please see, e.g., Upjohn Co. v. United States, 449 U.S. 383, 389 (1981) ("sound legal advice or advocacy serves public ends and . . . depends upon the lawyer's being fully informed by the client"); Trammel v. United States, 445 U.S. 40, 51 (1980) ("the advocate and counselor . . . [must] know all that relates to the client's reasons for seeking representation if the professional mission is to be carried out."); Hunt v. Blackburn, 128 U.S. 464, 470 (1888) (assistance of counsel "can only be safely and readily availed of when free from the consequences or the apprehension of disclosure"); *see also* Geoffrey Hazard, *An Historical Perspective on the Attorney-Client Privilege*, 66 CALIF. L. REV. 1061, 1061 (1978) (noting that the privilege allows an attorney to prepare a case and effectively advocate).

26. Privacy by design is the notion that privacy should be part of the development process of new technologies rather than tacked on at the end. For two comprehensive approaches to defining privacy by design, please see HARTZOG, PRIVACY'S BLUEPRINT, *supra* note 21; *see also* Ari Ezra Waldman, *Privacy's Law of Design*, 9 U.C. IRVINE L. REV. 1239 (2019) [hereinafter Waldman, *Privacy's Law of Design*].

There is no perfectly safe space, digital or otherwise. Even better design, comprehensive federal and state laws, and private ordering cannot account for all human mischief. But in a modern world in which sharing is, if not always mandatory, expected, law and design can make social spaces safer by supporting and protecting trust and repairing it when it breaks down.

I. TRUST AND SOCIAL GOVERNANCE

Although scholars bring different modalities to the study of trust, there is remarkable overlap in the way different fields conceptualize the concept. That literature has been discussed in depth elsewhere.[27] A chief take away from that scholarship is that trust is an essential element of online social governance. Joel Reidenberg[28] and Lawrence Lessig[29] predicted this when they argued that law, architecture, markets, and norms work together to regulate online conduct. Trust is one of those norms and, therefore, an important focal point for the study of technosocial spaces.

A. What is Trust?

Robert Putnam and Francis Fukuyama think about trust as epiphenomenal with social capital. For Putnam, social capital is a "feature of social organizations . . . that facilitates coordination and cooperation for mutual benefit."[30] Fukuyama goes a step further, arguing that social capital consists of norms or values, "instantiated in an actual relationship among two or more people, that promote cooperation between them."[31] On a micro level, social capital constitutes the advantages and benefits that individuals realize owing to their connections with others, like coworkers learning from one another and cooperating to achieve a goal[32] or social groups from diverse backgrounds whose experiences are enhanced because of their diversity.[33] Social capital also develops on a more macro level, among

27. *See* WALDMAN. PRIVACY AS TRUST, *supra* note 24, at 51–60; *see also* Richards & Hartzog, *supra* note 17.

28. *See* Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553, 554–55 (1998) ("The creation and implementation of information policy are embedded in network designs and standards as well as in system configurations.").

29. *See* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE, VERSION 2.0 (2006).

30. Robert D. Putnam, *Bowling Alone: America's Declining Social Capital*, 6 J. DEMOCRACY 65, 67 (1995); *see also* ROBERT D. PUTNAM, BOWLING ALONE: THE COLLAPSE AND REVIVAL OF AMERICAN COMMUNITY 19 (2000).

31. Fukuyama, *Differing Disciplinary Perspectives*, *supra* note 22, at 480.

32. *See, e.g.*, Michael Useem & Jerome Karabel, *Pathways to Top Corporate Management*, 51 AM. SOC. REV. 184 (1986).

33. *See, e.g.*, Ronald S. Burt, *The Contingent Value of Social Capital*, 42 ADMIN. SCI. Q. 339 (1997).

individuals in larger communities and nations and even among nations and peoples.[34] In all cases, social capital refers to the good things that develop out of our connections to others.

Trust is one of those good things. Trust is a resource of social capital concerning the expectations that others will behave according to accepted norms.[35] It is the "favorable expectation regarding other people's actions and intentions,"[36] or the belief that others will behave in a predictable manner. For example, if I ask a friend to hold my spare set of keys, I trust she will not break in and steal from me. When an individual speaks with relative strangers in a support group like AA, she trusts that they will not divulge her secrets.[37] Trust, therefore, includes a willingness to accept some risk and vulnerability toward others and steps in to grease the wheels of social activity in the absence of perfect knowledge: I cannot know for certain that my neighbor will not abuse her key privileges or that my fellow support group members will keep my confidences. As Niklas Luhmann has stated, trust begins where knowledge ends.[38] As such, trust allows me to interact with and rely on others.

Trust is essential online. Nearly two decades ago, Helen Nissenbaum presciently noted that trust is "key to the promise the online world holds for great and diverse benefits to humanity," including richer communities, engaged politics, and robust commerce, because "[p]eople shy away from territories they distrust."[39] That is just as true today. Corporate executives talk about ensuring a steady stream of customer data by gaining user trust and confidence.[40] Apple asks us if we "Trust this browser?" when we log in to iCloud on a new device. In 2013, Facebook conducted a study of its users to determine "how trustworthy" they think Facebook is overall.[41] The

34. *See* FRANCIS FUKUYAMA, TRUST: THE SOCIAL VIRTUES AND THE CREATION OF PROSPERITY (1995).

35. *See* Alejandro Portes & Julia Sensenbrenner, *Embeddedness and Immigration: Notes on the Social Determinants of Economic Action*, 98 AM. J. SOC. 1320, 1332 (1993).

36. Guido Möllering, *The Nature of Trust: From Georg Simmel to a Theory of Expectation, Interpretation and Suspension*, 35 SOC. 403, 404 (2001); *see also* Ken Newton & Sonja Zmerli, *Three Forms of Trust and Their Association*, 3 EUR. POL. SCI. REV. 169, 171 (2011); J. David Lewis & Andrew Weigert, *Trust as Social Reality*, 63 SOCIAL FORCES 967, 968 (1985).

37. *See* ALCOHOLICS ANONYMOUS, UNDERSTANDING ANONYMITY (2018), https://www.aa.org/pages/en_US/understanding-anonymity [https://perma.cc/8A8D-HLG3] [hereinafter UNDERSTANDING ANONYMITY].

38. *See* NIKLAS LUHMANN, TRUST AND POWER 5 (Howard Davies, John Raffan & Kathryn Rooney trans., Tom Burns & Gianfranco Poggi eds., 2017) (1979).

39. Helen Nissenbaum, *Securing Trust Online: Wisdom or Oxymoron*, 81 B.U. L. REV. 635, 636 (2001).

40. *See* Timothy Morey et al., *Customer Data: Designing for Transparency and Trust*, HARV. BUS. REV., May 2015, at 96.

41. *See* Brian Fung, *Facebook Wants to Know If Your Trust It. But It's Keeping All the Answers to Itself*, WASH. POST (Dec. 31, 2013), https://www.washingtonpost.com/news/the-switch/wp/2013/12/

Federal Trade Commission (FTC)[42] and the California Attorney General's Office[43] couch their recommendations for transparency in corporate data use as a way of inspiring consumer trust. And the work of scholars like Kirsten Martin shows that failing to meet the privacy expectations of users negatively impacts the trust those users have in the website and reduces user willingness to share.[44]

B. How Does Trust Develop?

Trust, like other norms of social life, can develop hierarchically from above. For example, legal rules can influence norms of behavior through the law's expressive power,[45] as when the Supreme Court declares race-based discrimination illegal and, over time, the illegality of discrimination is accepted as a moral imperative.[46] Fiduciary laws, medical malpractice law, and legally enforced canons of ethics are just three of the myriad rules and private ordering schemes that support trust norms from above.

Norms can also be influenced by design, as when a builder puts spikes on first floor window sills to prevent loitering and thereby influences community norms about private property, community, and vagrancy.[47]

31/facebook-wants-to-know-if-you-trust-it-but-its-keeping-all-the-answers-to-itself/ [https://perma.cc/43PA-83CU].

42. *See* FED. TRADE COMM'N, MOBILE PRIVACY DISCLOSURES: BUILDING TRUST THROUGH TRANSPARENCY 3–4 (2013), https://www.ftc.gov/sites/default/files/documents/reports/mobile-privacy-disclosures-building-trust-through-transparency-federal-trade-commission-staff-report/130201mobileprivacyreport.pdf [https://perma.cc/43PA-83CU]; *see also Protecting Mobile Privacy: Your Smartphones, Tablets, Cell Phones and Your Privacy, Hearing Before the Subcomm. for Privacy, Tech. & the Law of the S. Comm. on the Judiciary*, 112th Cong. 90 (2011) (statement of Alan Davidson, Director of Public Policy, Google, Inc.).

43. *See* CAL. DEP'T OF JUSTICE, MAKING YOUR PRIVACY PRACTICES PUBLIC: RECOMMENDATIONS ON DEVELOPING A MEANINGFUL PRIVACY POLICY 4 (2014), https://oag.ca.gov/sites/all/files/agweb/pdfs/cybersecurity/making_your_privacy_practices_public.pdf [https://perma.cc/2LNN-KRHQ].

44. See Kirsten Martin, *Understanding Privacy Online: Development of a Social Contract Approach to Privacy*, 137 J. BUS. ETHICS 551 (2016).

45. *See, e.g.*, Citron, *Law's Expressive Value*, *supra* note 9, at 407; Deborah Hellman, *The Expressive Dimension of Equal Protection*, 85 MINN. L. REV. 1, 3 n.10 (2000) (law is coercive and expressive of norms); Elizabeth S. Anderson & Richard M. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503, 1570–71 (2000) (what the law is establishes a set of agreed upon values).

46. *See* Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2043 (1996); *see also* CASS R. SUNSTEIN, ONE CASE AT A TIME: JUDICIAL MINIMALISM ON THE SUPREME COURT 158 (1999). This is not to say that norm generation through law's expressive power is perfect. Discrimination, white supremacy, and other odious forms of bigotry are still far too common today. Court decisions haven't changed that. *See* BRIAN LEVIN & JOHN DAVID REITZEL, CTR. FOR THE STUDY OF HATE & EXTREMISM, REPORT TO THE NATION: HATE CRIMES RISE IN U.S. CITIES AND COUNTIES IN TIME OF DIVISION & FOREIGN INTERFERENCE 3–4 (2018), https://csbs.csusb.edu/sites/csusb_csbs/files/2018%20Hate%20Final%20Report%205-14.pdf [https://perma.cc/DFM5-M6BL].

47. *See* UNPLEASANT DESIGN (Gordan Savicic & Selena Savic eds., 2013) (collecting and analyzing myriad common examples of how the design of mostly public spaces can deter antisocial

When technologies mediate social interaction, code plays this role, mandating or guiding norms of behavior online, [48] as when platforms automatically screen out content from sites like *InfoWars* or *Gateway Pundit* in an attempt to algorithmically combat conspiracy theories and fake news,[49] or when digital rights management prohibits reproduction of copyrighted material.[50]

Norms of trust can also emerge from below, through experience or explicit or implicit social cues. Experience gives us more data from which to judge the trustworthiness of others; keeping a friend's confidences for ten years gives them a stronger basis for trust than a single day. Explicit ("this is between us") and implicit cues (physically turning away from a crowd, huddling down, whispering) can also generate expectations of trust.[51] As can reciprocity, which establishes mutual vulnerability[52] and helps generate mutual feelings of cooperation and altruism.[53] Cues also allow us to trust strangers. For example, two people who share a stigmatizing social identity often create an instant bond of trust based on a shared set of narratives and experiences.[54] We are more willing to interact with others the more embedded they are in a familiar social network.[55] We tend to trust experts and professionals based on their degrees, transferring the trust we have in a school's reputation, which we know, to one of its graduates, whom we do

behavior, from uncomfortable benches and window sill spikes that discourage people from sitting or lying down to unflattering light that deters everything from congregation to intravenous drug use).

48. *See* LESSIG, CODE, *supra* note 22.

49. *See, e.g.*, Craig Silverman, *This Analysis Shows How Fake Election News Stories Outperformed Reals News on Facebook*, BUZZFEED NEWS (Nov. 16, 2016, 4:15 PM), https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook [https://perma.cc/M7DS-NA4S].

50. *See* Julie E. Cohen, *Cyberspace As/And Space*, 107 COLUM. L. REV. 210, 223–24 (2007) ("The Internet of digital rights management, take-down notices, and content filtering is a colony, in which permissible interactions are rigidly structured in the interest of an assertedly greater social good."); *see also* Cohen, *DRM and Privacy*, *supra* note 19, at 580–86 (describing how DRM are designed to constrain user behavior).

51. *See* ERVING GOFFMAN, BEHAVIOR IN PUBLIC PLACES 173 (1963) (noting how conversations in public closed to others may sometimes involve "huddling" down to keep confidences).

52. *See, e.g.*, Nancy R. Buchan, Rachel T.A. Croson & Robyn M. Dawes, *Swift Neighbors and Persistent Strangers: A Cross-Cultural Investigation of Trust and Reciprocity in Social Exchange*, 108 AM. J. SOC. 168, 170 (2002) (recognizing the vulnerability that emerges out of sharing information with others).

53. *See* Fukuyama, *Differing Disciplinary Perspectives*, *supra* note 22, at 491–93; *see also* WALDMAN, PRIVACY AS TRUST, *supra* note 24, at 53.

54. *See* Michele Williams, *In Whom We Trust: Group Membership as an Affective Context for Trust Development*, 26 ACAD. MGMT. REV. 377, 381–82, 385 (2001) (providing survey and experimental evidence of the power of associational membership to influence trust in others).

55. See, e.g., Mark Granovetter, *Economic Action and Social Structure: The Problem of Embeddedness*, 91 AM. J. SOC. 481, 490 (1985).

not.[56] And we often choose doctors based on recommendations from friends or colleagues.[57]

These mechanisms of trust generation operate online, as well. On dating apps, for example, users protect themselves by ensuring reciprocal sharing of personal information and images.[58] They listen to cues through extended text messaging.[59] Queer users tend to trust other queer users online because of a sense of shared struggle.[60] And we are willing to share more information on platforms like Facebook when we see more of our friends and intimates sharing, as well.[61] As James Grimmelmann has noted, all of our trusted friends "can't be wrong" that Facebook, or any other platform, is a safe place to interact.[62] Together, these endogenous forces foster organic trust and allow social interaction to occur.

II. SAFE SOCIAL SPACES

But trust cannot operate alone. Design and law must work together to buttress trust norms and ensure safe and socially beneficial disclosures. And we see it work all the time: AA meetings, corporate teams, and attorneys and their clients are just three examples. This Part describes trust-enhancing design and law in each.

Participants in AA meetings need to share information without fear that their condition, which unfortunately remains stigmatized,[63] will be publicized. Therefore, meetings are designed with mutually enforced strict confidentiality rules and operating structures that engender norms of trust. Privacy torts are also available in case something goes wrong.[64] Similarly,

56. *See* Patricia M. Doney et al., *Understanding the Influence of National Culture on the Development of Trust*, 23 ACAD. MGMT. REV. 601, 607 (1998) (discussing the role of transference of trust).

57. *See* Roni Caryn Rabin, *You Can Find Dr. Right, with Some Effort*, N.Y. TIMES (Sept. 29, 2008), www.nytimes.com/2008/09/30/health/30find.html [https://perma.cc/7H8N-9FSC].

58. *See* Waldman, *Law, Privacy, and Online Dating*, *supra* note 6.

59. *See id.*

60. *See id.*

61. *See* Ari Ezra Waldman, *Privacy, Sharing, and Trust: The Facebook Study*, 67 CASE W. RES. L. REV. 193 (2016) [hereinafter Waldman, *Privacy, Sharing, and Trust*]; *see also* Alessandro Acquisti, Leslie K. John & George Loewenstein, *The Impact of Relative Standards on the Propensity to Disclose*, 49 J. MARKETING RES. 160, 162 (2012) ("[W]hen people are surrounded by others who are revealing intimate details about their lives, they may conform to the prevailing norm of divulgence.").

62. James Grimmelmann, *Saving Facebook*, 94 IOWA L. REV. 1137, 1161 (2009).

63. *See, e.g.*, Georg Schomerus et al., *The Stigma of Alcohol Dependence Compared with Other Mental Disorders: A Review of Population Studies*, 46 ALCOHOL & ALCOHOLISM 105 (2011) (showing that alcohol-dependent persons tend to be viewed as far more responsible for their actions and provoke more social rejection and negative emotions from others than those living with substance-unrelated mental disorders); *see also* Robin Room, *Stigma, Social Inequality and Alcohol and Drug Use*, 24 DRUG & ALCOHOL REV. 143 (2005) (significant stigma associated with alcohol and drug abuse contributes to exclusion of those most in need of social support).

64. *See infra* Part II.A.

corporate teams working toward productivity goals must share ideas and learn from each other without having to worry that one member is going to decamp to a competitor (or a competing team within a company) with inside information. To mitigate that risk, workers deploy informal tactics to determine trust in team members on the ground, employers write in non-compete clauses into contracts, and the Computer Fraud and Abuse Act (CFAA) and the law of trade secrets limit what workers can take with them when they switch jobs.[65] Finally, individuals seeking counsel must be able to share details of their cases in order to obtain effective representation. Therefore, they rely on heuristics and personal recommendations to find lawyers they trust. And they do so in a context in which the legal profession itself has set up powerful ethical guidelines that carry punishments for noncompliance and in which malpractice law protects victims from negligent attorneys.[66] In all three of these case studies, endogenous design and exogenous law reinforce trust norms among social actors, keeping the contexts safe for intimate disclosures.

A. Alcoholics and Narcotics Anonymous

Disclosure is written into the DNA of AA. Individuals attend meetings to "share their experience, strength and hope with each other" in order to recover.[67] Meetings function with both narrative and discussion, where members introduce themselves and share stories about how alcohol has impacted their lives. After the narrative, other members may add their own perspective, share similar or related stories, or make suggestions on how to deal with an impulse to drink. Discussion ensues, or the meeting turns to other members who volunteer to introduce themselves as alcoholics and share their stories.[68] All AA members engage in "[m]eeting and talking and

65. *See infra* Part II.B.

66. *See infra* Part II.C.

67. ALCOHOLICS ANONYMOUS, THIS IS A.A.: AN INTRODUCTION TO THE A.A. RECOVERY PROBLEM 2 (2017) https://www.aa.org/assets/en_US/p-1_thisisaa1.pdf [https://perma.cc/C2XR-MTKT] [hereinafter THIS IS A.A.] It should be noted here that neither AA nor Narcotics Anonymous (NA) are perfect organizations. Its effectiveness is up for debate, and many criticize its inclusion of "god" in its literature and twelve steps. Given that anyone can attend a meeting as long as they admit their lack of control over alcohol consumption, AA can also be dangerous. *See, e.g.*, LANCE DODES & ZACHARY DODES, THE SOBER TRUTH: DEBUNKING THE BAD SCIENCE BEHIND 12-STEP PROGRAMS AND THE REHAB INDUSTRY (2014) (finding AA's success rate at just five to eight percent of people); *Twelve Steps to Dangers: How Alcoholics Anonymous Can Be a Playground for Violence-Prone Members*, PROPUBLICA (Nov. 29, 2014, 7:10 PM), https://www.propublica.org/article/how-alcoholics-anonymous-can-be-a-playground-for-violence [https://perma.cc/Q8MG-W2PE]. I take no position on the effectiveness of AA. The purpose of briefly profiling AA here is to show that it is a social space, with powerful disclosure norms that are backed by norms of trust.

68. *See, e.g.*, E.J. Khantzian & John E. Mack, *How AA Works and Why It's Important for Clinicians to Understand*, 11 J. SUBSTANCE ABUSE TREATMENT 77, 86–87 (1995) (discussing the importance of storytelling through case vignette).

helping other"[69] participants because only through sharing stories can members "observe[] and follow[] the successful experience[s]" of those who are sober.[70]

The long scholarly literature on AA[71] shows that disclosure is essential. Disclosure transforms participants into their own therapists.[72] It reinforces self-identification as having a substance abuse problem, a threshold step to recovery. And testifying about experiences brings new members into the fold by giving them models to follow and proving that they are not alone.[73]

Trust is what allows AA members to meet their disclosure obligations in safety. Trust organically develops in each AA meeting because all participants start by knowing exactly one thing about each other: they all share a stigmatizing identity. Familiarity has long been understood by social scientists as a basis for trust among persons. Max Weber thought that shared Protestantism allowed people who did not really know each other to trust that they would be competent contractual partners.[74] It signaled their common values. Sharing an out-group identity signals a common narrative and common struggle, as well, both of which influence values.[75] Everyone in AA shares a common and difficult relationship with alcohol. They

69. THIS IS A.A., *supra* note 67, at 8.

70. *Id.* at 13.

71. The literature on AA is largely about its effectiveness and, therefore, far beyond the scope of this Article. *See, e.g.*, Lee Ann Kaskutas, *Alcoholics Anonymous Effectiveness: Faith Meets Science*, 28 J. ADDICTIVE DISEASES 135 (2009) (reviewing the literature on AA effectiveness); Henry A. Montgomery, William R. Miller & Scott Tonigan, *Does Alcoholics Anonymous Involvement Predict Treatment Outcome*, 12 J. SUBSTANCE ABUSE TREATMENT 241, 245 (1995) (showing that attending AA meetings did not have a statistically significant effect on sobriety after inpatient treatment compared to non-attenders); Rudolf H. Moos & Bernice S. Moos, *Participation in Treatment and Alcoholics Anonymous: A 16-Year Follow-Up of Initially Untreated Individuals*, 62 J. CLINICAL PSYCH. 735, 745–46 (2006) (finding that individuals who participated in AA for twenty-seven weeks or more had better sixteen-year alcohol-related outcomes than those who did not participate in any AA meetings).

72. *See* Frank Reissman, *The 'Helper' Therapy Principle*, 10 SOC. WORK 27, 27–28 (1965) (noting that sharing one's story can provide both help to others and personal benefit).

73. Although there is fierce debate over the effectiveness of self-help programs, there is considerably less disagreement in the clinical psychology literature about the self-help benefits of self-disclosure, narrative, and testimony, whether in a clinician's office or elsewhere. *See, e.g.*, JOHN MCLEOD, NARRATIVE AND PSYCHOTHERAPY (1997); *see also* James W. Pennebaker & Janel D. Seagal, *Forming a Story: The Health Benefits of Narrative*, 55 J. CLINICAL PSYCH. 1243, 1243–44 (1999) (the act of writing down a personal narrative can contribute to positive mental health benefits); JoAnne Banks-Wallace, *Emancipatory Potential of Storytelling in a Group*, 30 J. NURSING SCHOLAR. 17, 17–18 (1998) (discussing the importance of narrative storytelling in advancing the mental health of African women).

74. *See* MAX WEBER, *The Protestant Sects and the Spirit of Capitalism*, *in* FROM MAX WEBER: ESSAYS IN SOCIOLOGY 302, 312 (H. H. Gerth & C. Wright Mills eds. & trans., 1946).

75. *See* Martin Tanis & Tom Postmes, *A Social Identity Approach to Trust: Interpersonal Perception, Group Membership and Trusting Behaviour*, 35 EUR. J. SOC. PSYCH. 413 (2005) (finding that when individuals are not identifiable, trusting behavior is based on expectations of reciprocity inferred from group membership); Williams, *supra* note 54, at 381, 385 (discussing how "[p]eople tend to associate positive beliefs and feelings with the groups to which they belong").

reciprocally share their stories, building the kind of mutual vulnerability that generates trust. As a result, when AA members notice an acquaintance at a meeting, AA advises them not to worry: AA expects that members will "respect your privacy."[76] AA literature also notes that "experience suggests that AA members . . . are vigilant" about protecting confidentiality, only use first names, and maintain personal anonymity in the media.[77] Members behave like this on their own because of the common struggle they share.

These organic norms of trust are supported by design. AA designs in trust by creating rules of confidentiality. Anonymity is the "spiritual foundation" of AA.[78] According to AA's founder, appropriately only known as Bill W., anonymity provides "protection for the newcomer, respect and support of the outside world, and security from those of us who would use A.A. for sick and selfish purposes."[79] The knowledge commons of AA are, therefore, governed by rules, guidelines, and suggestions that buttress the already strong confidentiality norms that develop immediately at AA meetings.[80]

And the confidentiality of AA meetings can also be buttressed by law. Although cases here are few and far between for obvious reasons—namely, filing suit for being exposed by another as a member of AA requires going public about the plaintiff's membership[81]—a few courts have supported the privacy expectations of AA members based on the norms of trust inherent in AA itself. In *State v. Ashworth*,[82] for example, police officers investigating a report of gun shots interrupted an AA meeting, pulled Ashworth outside, and conducted several field sobriety tests that put his blood-alcohol content above the legal limit.[83] Because the police lacked a warrant, Ashworth moved to suppress evidence of the alcohol concentration in his blood.[84] The trial court stated that Ashworth's expectation of privacy

76. *See* UNDERSTANDING ANONYMITY, *supra* note 37, at 8.

77. *Id.* at 12.

78. *See* ALCOHOLICS ANONYMOUS, A.A. GUIDELINES (2017), https://www.aa.org/assets/en_US/mg-18_internet.pdf [https://perma.cc/TF45-NTPM] [hereinafter A.A. GUIDELINES]; *Anonymity is the Spiritual Foundation*, NA WAY MAG., July 2002, at 6, https://www.na.org/admin/include/spaw2/uploads/pdf/naway/en/usnaway_jul2002.pdf [https://perma.cc/B5RJ-KNQJ].

79. A.A. GUIDELINES, *supra* note 78, at 1.

80. The term "knowledge commons" refers to information that is a shared resource. That the people in a given AA meeting are all alcoholics is an example of a piece of knowledge in common to everyone at that meeting. *See* GOVERNING MEDICAL KNOWLEDGE COMMONS (Katherine J. Strandburg, Michael J. Madison & Brett Frischmann eds., 2014).

81. It is, of course, possible to plead pseudonymously or anonymously in litigation. *See, e.g.*, Joan Steinman, *Public Trial, Pseudonymous Parties: When Should Litigants be Permitted to Keep Their Identities Confidential?*, 37 HASTING L.J. 1 (1985) (providing an analytical scheme for determining when pseudonymity in litigation should be allowed).

82. 228 P.3d 381 (Idaho 2010).

83. *Id.* at 382.

84. *Id.* at 382. Because Ashworth had a previous DUI conviction, his alcohol concentration of above 0.20 meant that he had committed a felony under Idaho law. *Id.* (citing IDAHO CODE § 18-8004C).

rested on "his and society's understanding of the privacy afforded by attendance at an AA meeting." The court explained:

> Certainly, if a group member admitted during a meeting that he had driven to the meeting while drunk, neither society nor members would consider that information public currency. The nature of Alcoholics Anonymous conveys an objective understanding of a group that protects the anonymity of its members. Information imparted to the group in a meeting where the privacy of those attending is expected is intended to be held in confidence by other members.[85]

The Idaho trial court in *Ashworth* was willing to stand behind and support that built-in trust norm.[86]

The *Ashworth* court is not alone in using the law to support the trust and confidentiality norms of AA meetings. In *Harford v. City of Santa Clarita*,[87] a town resident and nonsmoker petitioned the court to order the township police to enforce a municipal no-smoking ordinance outside a meeting place of the Rafters Group, the local chapter of AA.[88] The ordinance made smoking illegal in public places, but exempted so-called "private clubs" and gave city officials discretion to interpret and apply the law.[89] The Rafters argued that it fell within the exemption. After several notifications of smoking in violation of the ordinance, city officials declined to enforce the law, arguing that AA fell under the private club exception.[90] The appellate court agreed, noting that it was entirely reasonable for city officials to consider the privacy normally afforded to AA meetings to determine that it was a private club entitled to conduct its meetings pursuant to privacy and association rights guaranteed by the California Constitution.[91]

In the United Kingdom, where the law of confidence is far more developed than in the United States,[92] the law protects the trust that organically develops inside support groups. In one famous example, the

85. *Id.* at 385.

86. This motion to suppress was ultimately rejected by the intermediate appellate court, but not because it found the trial court wrong in its use of law to support AA's trust norms. Rather, it was a problem of evidence. Ashworth had not bothered to prove at trial that he had both a subjective and objective expectation of privacy in the AA meeting. The court, therefore, reversed the grant of the motion to suppress based on clear precedent requiring a high evidentiary showing from similar petitioners. *Id.* at 385–86.

87. No. B144255, 2002 WL 27113 (Cal. Ct. App. Jan. 10, 2002).

88. *Id.* at *1.

89. *Id.* at *5.

90. *Id.* at *1.

91. *Id.* at *5–11.

92. *See* Neil M. Richards & Daniel J. Solove, *Privacy's Other Path: Recovering the Law of Confidentiality*, 96 GEO. L.J. 123, 145–73 (2007).

supermodel Naomi Campbell sued the British tabloid, *The Mirror*, after it published a photograph of her leaving a Narcotics Anonymous (NA) meeting.[93] Although she admitted the public may have an interest in news about her, Ms. Campbell nevertheless argued that publishing the picture was an invasion of the privacy that she should expect as a member of NA.[94] The House of Lords agreed. Lord Hope wrote that "the details of Miss Campbell's attendance at Narcotics Anonymous [were] private information which imported a duty of confidence."[95] Lady Hale, also in the majority, called both the fact of NA attendance and any information disclosed in NA meetings "obviously private" and "both private and confidential, because it related to an important aspect of Miss Campbell's physical and mental health."[96] To hold otherwise would risk Ms. Campbell's and others' mental health by discouraging future attendance at NA or AA meetings.[97] Although the relative weakness of the tort of breach of confidentiality[98] and the enormous newsworthiness exception to the tort of public disclosure of private facts[99] would make Ms. Campbell's case more difficult in the United States, it is clear that law can and has supported the organic and privately ordered trust norms that make programs like AA and NA work.

B. Teams of Coworkers

The information exchanged among coworkers may be different than the kind of information shared among AA members. But information sharing is no less essential to achieving the team's socially beneficial productivity goals. Teams are, after all, the fundamental work unit of most

93. Campbell v. MGN Ltd. [2004] UKHL 22, [2004] 2 AC 457 (Eng.).

94. *Id.* at [2]. There was some disagreement among the Lords about the proper claim for the case. Suffice it to say, regardless of which tort was operative in this case, the House of Lords was willing to recognize that British common law can recognize and support the confidentiality norms embedded in AA as legitimate expectations of privacy.

95. *Id.* at [95].

96. *Id.* at [147].

97. *Id.* at [155].

98. *But see* Ari Ezra Waldman, *A Breach of Trust: Fighting "Revenge Porn"*, 102 IOWA L. REV. 709, 722–28 (2016) (arguing that there no doctrinal barriers to the expansion of the breach of confidentiality tort in the United States).

99. *See* RESTATEMENT (SECOND) OF TORTS, § 652D. The newsworthiness exception has been broadly interpreted, making it difficult for both well-known and even private individuals to recover under the tort of public disclosure. *See, e.g.*, Haynes v. Alfred A. Knopf, Inc., 8 F.3d 1222, 1232 (7th Cir. 1993) ("[T]he First Amendment greatly circumscribes the right even of a private figure to obtain damages for the publication of newsworthy facts about him, even when they are facts of a kind that people want very much to conceal."); Sidis v. F-R Publ'g Corp., 113 F.2d 806, 809 (2d Cir. 1940) (holding that intimate details of a former public figure's private life are not entitled to kept out of the press).

organizations.[100] Teams of coworkers make hiring decisions in midsized and large organizations; engineers work together to build technology products; juries decide guilt. Scholars have found increased knowledge transfer within and among teams in organizations correlate with complex problem solving,[101] better productivity and performance,[102] and organizational learning,[103] or company-wide adaptions to new realities.[104] These benefits are based on the capacity of discussion and sharing to increase the pool of available knowledge, allowing teams to make theoretically better, more accurate, and more reflective decisions.[105]

Norms of disclosure are, therefore, built into teams. Discussion is the bread and butter of team function, with discussion structures varying widely.[106] Team members are expected to share stories and insights during work,[107] provide feedback,[108] and discuss together how best to solve complex problems.[109] One study of software and hardware design teams showed that information sharing is an essential part of the design process. Design teams, which are usually constituted by individuals or subgroups working on specific parts of the larger product, often hold weekly meetings

100. *See* PETER M. SENGE, THE FIFTH DISCIPLINE: THE ART & PRACTICE OF THE LEARNING ORGANIZATION (2006); Paul Osterman, *How Common is Workplace Transformation and Who Adopts It?*, 47 INDUS. LAB. RELATIONS REV. 172 (1994).

101. *See, e.g.*, Lazer & Friedman, *supra* note 25, at 668–69.

102. *See, e.g.*, Jessica R. Mesmer-Magnus & Leslie A. DeChurch*, Information Sharing and Team Performance: A Meta-Analysis*, 94 J. APPLIED PSYCH. 535 (2009) (reviewing 72 independent studies, covering 4,795 total groups with 17,279 persons, and finding that information sharing positively predicted team performance across all levels); Morten T. Hansen, *Knowledge Networks: Explaining Effective Knowledge Sharing in Multiunit Companies*, 13 ORG. SCI. 232 (2002) (finding that increased knowledge sharing among company units contributed to faster completion of projects).

103. *See, e.g.*, Edmondson, *Learning in Organizations*, *supra* note 25, at 633–37 (linking team learning to organizational learning and adaptation).

104. *Id.* at 631 (defining organizational learning as "a process of improving organizational actions through better knowledge and understanding").

105. Notably, this does not always work. In a famous study conducted by the psychologists Garold Stasser and William Titus in 1985, the researchers showed that group discussion is often biased toward already shared or commonly known information, rather than the new information from which teams could benefit. Stasser and Titus found, then, that a team can confirm previously held views rather than open team members to new ideas. *See* Garold Stasser & William Titus, *Pooling of Unshared Information in Group Decision Making: Biased Information Sampling During Discussion*, 48 J. PERSONALITY & SOC. PSYCH. 1467, 1467–68 (1985). The balance of research still suggests that knowledge and information sharing in teams can be successful and is positively correlated with various positive social outcomes. *See* Mesmer-Magnus & DeChurch, *supra* note 102, at 535–41 (reviewing the literature after 1985).

106. *See* Gwen M. Wittenbaum & Jonathan M. Bowman, *A Social Validation Explanation for Mutual Enhancement*, 40 J. EXPER. SOC. PSYCH. 169 (2004) (discussing how different models of discussion, including the role of positive feedback, affect information gathering in teams).

107. *See* John Seely Brown & Paul Duquid, *Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation*, 2 ORG. SCI. 40, 40–41 (1991).

108. *See* Teresa K. Lant, *Aspiration Level Adaptation: An Empirical Exploration*, 38 MGMT. SCI. 623, 624 (1992) (finding that teams respond to performance feedback).

109. *See* Lazer & Friedman, *supra* note 25, at 668–69.

where they learn about each other's progress. Both leaders and team members expect that everyone will share their work, new information that impacts the team's responsibilities, and questions that have come up during design.[110] And questions always come up; new information, whether from other teams, suppliers, managers, mentors, or scholars, is always needed to complete the design phase.[111] Therefore, team members work together to identify the best sources of information, collaborate with others in the company, and conduct outside research. Primarily, though, team members talk to each other: diverse team members often have answers or recommendations on where to turn. Serendipitous encounters outside offices or in hallways also help share information, and project managers are usually expected to mediate and find connections that help fill information gaps.[112]

Just like sharing in AA meetings is necessary for recovery, sharing in teams is necessary for productivity. And norms of trust are essential to sustain both types of sharing. Trust gives coworkers the comfort and safety to share ideas freely,[113] to reflect on performance,[114] and to suggest new ways of doing business without fear of opportunistic behavior from would-be competitors.[115]

Also like AA, where trust organically exists in testimonial meetings because each member shares a stigmatizing social identity,[116] some level of trust exists organically in corporate teams by virtue of their shared membership in the organization and shared goal of production. The management scholars Wenpin Tsai and Sumantra Ghoshal called this phenomenon the result of a "shared vision" that helps a team or network direct their efforts.[117] Having a shared vision, common membership, and identical goals bonds teammates together, creating a foundation of trust.[118] Trust also organically develops within teams that exhibit stable membership over time. Learning depends, at least in some part, on memories of individuals: a team member can only advise against repeating mistakes if

110. *See* Steven Poltrock et al., *Information Seeking and Sharing in Design Teams*, 2003 INT. ACM SIGGROUP CONF. ON SUPPORTING GRP. WORK 239, 240–41.

111. *Id.* at 242.

112. *Id.* at 244.

113. Amy Edmondson, *Psychological Safety and Learning Behavior in Work Teams*, 44 ADMIN. SCI. Q. 350, 350 (1999) [hereinafter Edmondson, Psychological Safety and Learning Behavior].

114. *See* Edmondson, *Learning in Organizations supra* note 25, at 633–34.

115. *See, e.g.*, Andrew C. Inkpen & Eric W. K. Tsang, *Social Capital, Networks, and Knowledge Transfer*, 30 ACAD. MGMT. REV. 146, 154 (2005); J. Carlos Jarillo, *On Strategic Networks*, 9 STRATEGIC MGMT. J. 31, 37 (1988).

116. *See* text accompanying notes 74–77.

117. *See* Wenpin Tsai & Sumatra Ghoshal, *Social Structure of "Cooperation" Within a Multiunit Organization: Coordination, Competition, and Intraorganizational Knowledge Sharing*, 13 ORG. SCI. 179 (2002).

118. *See* Inkpen & Tsang, *supra* note 115, at 159.

she was around for the mistakes in the first place. Moreover, maintaining stability in teams allows coworkers to develop interpersonal relationships and have the kind of repeated social exchanges that build trust.[119]

Because of the importance of trust within teams, companies design in trust mechanisms. Managers schedule in-person meetings to both ensure face-to-face relationship-building and to have some control over the flow of information.[120] To build trust among team members and incent productivity, corporations create internal awards to highlight successful employees. Although competitions for awards can create intra-team competition, having clear and transparent award criteria, honoring more than one worker at a time, and rejecting zero-sum awards can increase trust within teams by assuring other team members that a coworker will be able to pull her weight.[121]

Companies also embed non-disclosure and non-compete clauses into employment contracts to support what Amy Edmondson calls the "psychological safety" of trust within teams.[122] The enforceability of covenants not to compete vary from state to state.[123] But they can reassure employees that it is difficult for their coworkers to steal a good idea shared during a design meeting and engage in opportunistic behavior. Knowing that everyone is subject to the same limitation, coworkers can feel free to disclose new ways of solving problems without fear of losing a competitive advantage in the marketplace.[124]

---

119. *See id.* at 156; *see also* Kathleen Carley, *Organizational Learning and Personnel Turnover*, 3 ORG. SCI. 20, 22 (1992) (noting that memories of individuals on teams is essential for team and organizational learning).

120. *See* Poltrock et al., *supra* note 110, at 242.

121. *See* Inkpen & Tsang, *supra* note 115, at 158.

122. Edmondson, *Psychological Safety and Learning Behavior*, *supra* note 113, at 350. *But see* THOMAS O. MCGARITY & WENDY E. WAGNER, BENDING SCIENCE: HOW SPECIAL INTERESTS CORRUPT PUBLIC HEALTH RESEARCH 111 (2008) (arguing that "companies can invoke confidentiality clauses in employee contracts" to prevent current and former employees from revealing "secret research").

123. Different states approach covenants not to compete differently. In 2016, President Obama issued an executive order that called for state legislatures to reform non-compete law. Exec. Order No. 13,725, 81 Fed. Reg. 23,417 (Apr. 15, 2016). A related "Call to Action" urged states to (1) ban noncompete clauses for low-earning workers; (2) improve transparency and fairness; and (3) incentize employers to write enforceable contracts. WHITE HOUSE, STATE CALL TO ACTION ON NON-COMPETE AGREEMENTS, https://obamawhitehouse.archives.gov/sites/default/files/competition/ noncompetes-calltoaction-final.pdf [https://perma.cc/PCM4-6W5J]. California generally rejects all non-compete clauses. *See* CAL. BUS. & PROF. CODE § 16600 (2017). So does Colorado. *See* COL. REV. STAT. § 8-2-113(2) (2017). Illinois strictly applies a reasonableness requirement. *See, e.g.*, Cambridge Eng'g, Inc. v. Mercury Partners 90 BI, Inc., 879 N.E.2d 512 (Ill. 2007). Washington State is more permissive. *See, e.g.*, Perry v. Moran, 748 P.2d 224, 229–30 (Wash. 1987); Knight, Vale & Gregory v. McDaniel, 680 P.2d 448, 451–52 (Wash. Ct. App. 1984). So is Massachusetts. *See* Marcam Corp. v. Orchard, 885 F. Supp. 294, 299 (D. Mass. 1995).

124. *See* Inkpen & Tsang, *supra* note 115, at 158 (when members worry that they may be competing against each other, suspicion replaces trust and knowledge sharing is sacrificed).

The law outside the company serves the same function. Many state courts are willing to enforce non-compete clauses and non-disclosure agreements.[125] The Computer Fraud and Abuse Act (CFAA) allows employers to sue employees who "intentionally [access] a computer without authorization or [exceed] authorized access" to steal company secrets.[126] And the law of trade secrecy prohibits employees from transferring corporate proprietary information to gain an unfair advantage.[127] Trade secret law is overtly utilitarian in this respect: one of its primary purposes is to give companies—and, by extension, their employees doing the work—the breathing space to innovate, create relationships, and share information knowing that bad, opportunistic behavior will be punished.[128] This breathing space for innovation is another way of understanding the role of trust among teams of coworkers, and it only survives because it is buttressed by design and backed by law.[129]

125. In addition to the cases cited in note 123, other cases show many states are willing to enforce non-compete clauses. *See, e.g.*, Vital Images, Inc. v. Martel, No. 07-4195, 2007 WL 3095378, at *3 (D. Minn. Oct. 19, 2007); Borg-Warner Protective Servs. Corp. v. Guardsmark, Inc., 946 F. Supp. 495, 501 n.6 (E.D. Ky. 1996); Raimonde v. Van Vlerah, 325 N.E.2d 544, 547 (Ohio 1975); Allen v. Rose Park Pharmacy, 237 P.2d 823, 826 (Utah 1951).

126. *See* 18 U.S.C. § 1030(a)(2) (2008); *see also* United States v. John, 597 F.3d 263 (5th Cir. 2010); United States v. Rodriguez, 628 F.3d 1258, 1263 (11th Cir. 2010); Int'l Airport Centers, LLC v. Citrin, 440 F.3d 418 (7th Cir. 2006); EF Cultural Travel BV v. Explorica, Inc., 274 F.3d 577 (1st Cir. 2001).

127. *See* Harlan M. Blake, *Employee Agreements Not to Compete*, 73 HARV. L. REV. 625, 673–74 (1960); *see also* RESTATEMENT (SECOND) OF TORTS, § 757, Comment b (2017).

128. *See, e.g.*, Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 481–82 (1974) (discussing the purposes behind trade secret law); Wexler v. Greenberg, 160 A.2d 430, 434–35 (Pa. 1960) (focusing on the importance of trade secrecy in research and development).

129. Undoubtedly, legal levers like covenants not to compete and trade secrecy can have negative social effects—the former are often thrust upon employees in contexts of unequal bargaining power, and companies use the latter to shield themselves from legal and public scrutiny. Rebecca Wexler, for example, has shown that trade secrecy is used in the criminal justice system to hide the ways algorithms decide the fate of some convicted of crimes. *See* Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343 (2018) (showing how trade secrecy is invoked in the criminal justice system). And according to Rachel Arnow-Richman, unequal bargaining positions make it "inappropriate to view noncompete terms as the product of reasoned reflection or as dispositive of the parties' rights and obligations." Rachel S. Arnow-Richman, *Bargaining for Loyalty in the Information Age: A Reconsideration of the Role of Substantive Fairness in Enforcing Employee Noncompetes*, 80 OR. L. REV. 1163, 1215 (2001). This is true even for employees with particularly valuable skills: "[e]ven if a particular employee possesses valuable human capital that is in demand in the relevant market, . . . there are [substantive and procedural] reasons to distrust the quality of the bargain he or she reaches with the employer." *Id.* at 1214.

How, or whether, to deploy these tools are important sociolegal questions beyond the scope of this paper. For now, it is sufficient to note that one of the goals of covenants not to compete and trade secrecy is to create a zone of confidentiality around information sharing in the workplace. In that capacity, they support organic and designed-in norms of trust in order to foster socially beneficial disclosure.

C. Attorney-Client Relationships

The attorney-client relationship is, like AA meetings and corporate teams, an overtly social environment where disclosure is both necessary and socially beneficial. As such, disclosure is designed into the relationship itself. The Supreme Court has recognized this, noting in *UpJohn v. United States*[130] that effective advocacy depends on the "lawyer's being fully informed by the client."[131]

The Supreme Court also long ago acknowledged the salient role of trust in the attorney-client relationship, noting in 1888 that a lawyer's "assistance can only be safely and readily availed of when free from the consequences or the apprehension of disclosure."[132] Therefore, trust norms compensate for the vulnerability inherent in sharing personal information with an attorney.[133] Just like in AA meetings and within teams, that trust begins organically. We trust experts and other professionals based on their degrees.[134] There is some evidence that we trust lawyers and doctors based on firm or hospital affiliations, respectively,[135] and even office design.[136] The transference process does not end there. Many of us do not choose doctors and lawyers based solely on their degrees. Rather, we rely on the recommendations of others and, in particular, those that we respect.[137]

But because going to a good law school or even coming recommended does not guarantee a lawyer will always act in her client's interests,[138] the legal profession designs in its own rules that support the organic trust of an attorney-client relationship. That private ordering comes from three sources: the oath that every lawyer takes upon being sworn into the bar;[139]

130. 449 U.S. 383 (1981).

131. *Id.* at 389.

132. Hunt v. Blackburn, 128 U.S. 464, 470 (1888); *see also* Trammel v. United States, 445 U.S. 40, 51 (1980).

133. *See* CHARLES W. WOLFRAM, MODERN LEGAL ETHICS 146 (1986) (noting that trust is important in attorney-client relationships to ensure full and frank disclosure).

134. Doney et al., supra note 56, at 603.

135. Mark A. Hall et al., *Trust in Physicians and Medical Institutions: What Is It, Can It Be Measured, and Does It Matter?*, 79 MILBANK Q. 613, 619–20 (2001).

136. *The Best Law Firm Offices in America: The Finalists!*, ABOVE THE LAW (Aug. 30, 2012, 6:19 PM), http://abovethelaw.com/2012/08/the-best-law-firm-offices-in-america-the-finalists/2 [https://perma.cc/K2LB-3R 28].

137. Rabin, *supra* note 57.

138. Acting in a client's interests is the *sin qua non* of fiduciary law. Fiduciaries are those that have special obligations of loyalty to another. Those loyalties are based on trust: a trustor, client, or beneficiary hands over money, control, and information to another, who, in turn, has a duty not to betray that trust. *See* TAMAR FRANKEL, FIDUCIARY LAW 106–08 (2011).

139. The oath of admission for the United States federal courts, for example, is as follows: "I _____ do solemnly swear (or affirm) that as an attorney and as a counselor of this court I will conduct myself uprightly and according to law, and that I will support the Constitution of the United

the American Bar Association, which has developed Model Rules of Professional Conduct;[140] and bar association ethics opinions, which interpret the rules and give lawyers guidance on particular ethical dilemmas.[141] At the heart of these ethical codes and guidelines is confidentiality.[142] Together, these self-regulatory tools remind clients that they can speak freely with their attorneys, share personal, even stigmatizing information, trusting that their lawyer will keep their confidences.

That trust is also enforced by the courts and other legal levers. Ethics rules are enforced by state disciplinary regimes.[143] Courts also rely on codes of professional ethics when poorly-served clients bring motions to disqualify.[144] The law of professional malpractice allows clients to sue their attorneys for, among other things, betraying their confidences.[145] And the attorney-client privilege protects communications between lawyers and their clients in court. In *Trammel v. United States*,[146] the Court stated that the rationale for this, and the priest-penitent and doctor-patient, privileges is explicitly based on trust:

> These privileges are rooted in the imperative need for confidence and trust. The priest-penitent privilege recognizes the human need to disclose to a spiritual counselor, in total and absolute confidence, what are believed to be flawed acts or thoughts and to receive priestly consolation and guidance in return. The lawyer-client privilege rests on the need for the advocate and counselor to know all that relates to the client's reasons for seeking representation if the professional mission is to be carried out. Similarly, the physician must know all that a patient can articulate in order to identify and to treat disease;

States." *See Attorney Oath of Admission*, U.S. COURTS, http://www.uscourts.gov/sites/default/files/ao153_0.pdf [https://perma.cc/P3SR-NUWM].

140. *See* MODEL RULES OF PROF'L CONDUCT (AM. BAR ASS'N 2016) [hereinafter MODEL RULES], https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/model_rules_of_professional_conduct_table_of_contents/ [https://perma.cc/65FP-9CMU].

141. *See* Peter A. Joy, *Making Ethics Opinions Meaningful: Toward More Effective Regulation of Lawyers' Conduct*, 15 GEO. J. LEG. ETHICS 313 (2002) (discussing the role of ethics opinions).

142. *See, e.g.*, MODEL RULES r. 1.6, https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/rule_1_6_confidentiality_of_information/ [https://perma.cc/8TG3-KESK].

143. *See* Leslie C. Levin, *The Emperor's Clothes and Other Tales About the Standards for Imposing Lawyer Discipline Sanctions*, 48 AM. U. L. REV. 1, 1–4 (1998) (noting that since the ABA's Special Committee on Evaluation of Disciplinary Enforcement chastised lawyer discipline as "scandalous," the role of the state in enforcing attorney ethical standard has increased).

144. *See, e.g.*, Silver Chrysler Plymouth, Inc. v. Chrysler Motors Corp., 518 F.2d 751, 753 (1975) (starting with attorney ethics rules during a motion to disqualify).

145. Benjamin C. Zipursky, *Legal Malpractice and the Structure of Negligence Law*, 67 FORDHAM L. REV. 649, 662–82 (1998) (describing legal malpractice as a subset of negligence law and identifying gaps in the model for tortious legal malpractice developed by William Prosser).

146. 445 U.S. 40 (1980).

barriers to full disclosure would impair diagnosis and treatment.[147]

Therefore, law, through the use of the courts as enforcement mechanisms and the doctrines of malpractice and privilege, helps balance out the powerful disclosure norms in attorney-client relationships with equally as powerful norms of trust.

III. THE PROBLEM OF TECHNOSOCIAL SPACES

No social space is perfectly safe; there is no way to guard against all opportunistic or malicious behavior. AA members are sometimes outed, workers steal secrets, and lawyers betray confidences. But within those social spaces, actors share with the expectation, designed in and backed by law, that their information will be kept confidential. And society benefits as a result. That's how it's supposed to work.

Technosocial spaces, or those in which technology mediates social interaction, are different. Although they, like offline contexts, are characterized by powerful pressures to disclose, their technical architecture, internal rules, and the legal and regulatory environment in which they operate do not support the kind of trust that protects users from harm. In fact, they do the opposite. Technosocial platforms are designed to entice and manipulate disclosure with false trust. And the law lets them do it. These missing pieces make these spaces unsafe, and ripe for harassment and invasions of privacy. This is evident from two case studies described below: online social networks and interactions with social robots.

A. Disclosure and Other Risks

Disclosure is the lifeblood of technosocial spaces. Without sharing our likes, opinions, and behaviors, technosocial platforms could not achieve their goals of bringing people together.[148] Nor could they learn from us, adapt to our needs, and provide the kinds of conveniences consumers seem to want. And they certainly couldn't process, analyze, and sell our data for profit.[149] Facebook needs our data to sell billions of dollars in targeted advertising space. Artificial intelligence needs our data to learn. Dating apps

147. *Id.* at 51. The phrase "confidence and trust" as a basis for the privilege doctrine is repeated in ninety-eight federal, state, and international cases citing *Trammel*.

148. Facebook's mantra is "bring the world closer together." *See* Josh Constine, *Facebook Changes Mission Statement to 'Bring the World Closer Together'*, TECHCRUNCH (June 26, 2017), https://techcrunch.com/ 2017/06/22/bring-the-world-closer-together/ [https://perma.cc/5ETB-PFH8].

149. *See, e.g.*, Mark Hachman, *The Price of Free: How Apple, Facebook, Microsoft and Google Sell You to Advertisers*, PCWORLD (Oct. 1, 2015, 3:00 AM), https://www.pcworld.com/article/2986988/privacy/the-price-of-free-how-apple-facebook-microsoft-and-google-sell-you-to-advertisers.html [https://perma.cc/4LBX-K7D8].

need our data to match us to others. Disclosure is a necessary part of technosocial spaces.

This poses risks to our safety. Twitter has a long history of tolerating hate and harassment on its platform, contributing to the silencing of women and users from other marginalized communities.[150] Some dating apps and websites do little even when they know their platforms are rife with racism, homophobia, and harassment.[151] And some websites invite their users to post images of nonconsensual pornography.[152] Our privacy is also in danger. Facebook's has for years taken a cavalier approach to third-party access to user data, giving companies we don't know access to our personal information.[153] Google, like a myriad of other digital platforms, mines terabytes of personal data for behavioral targeting.[154] Snapchat made us think we had control over the shelf lives of the photos and videos we sent to other users, but the reality was far different.[155] The list goes on.

Social robots,[156] like PARO, the therapeutic baby seal,[157] or Sony's Aibo dog.[158] They pose some similar and some unique dangers by virtue of their social abilities, including communication, cooperation, and learning.[159] Social robots are machines that collect vast amounts of data behind a veil of human-like social features.[160] They can also perpetuate intimate partner

150. *See, e.g.*, Monique Judge, *Twitter Has a Serious Harassment and Abuse Problem but Doesn't Seem to Want to Cure It*, ROOT (Oct. 30, 2017, 4:59 PM), https://www.theroot.com/twitter-has-a-serious-harassment-and-abuse-problem-but-1819979725 [https://perma.cc/D65X-F96A]; *see also* Citron, *Cyber Civil Rights*, *supra* note 3.

151. *See, e.g.*, Herrick v. Grindr, LLC, 306 F. Supp. 3d 579 (S.D.N.Y. 2018); *see also* Waldman, *Law, Privacy, and Online Dating*, *supra* note 6 (discussing the experience of some users who report notifying platforms of terms of service violations, but never receiving responses or remediation).

152. *See* Danielle Keats Citron & Mary Anne Franks, *Criminalizing Revenge Porn*, 49 WAKE FOREST L. REV. 343 (2014).

153. *See, e.g.*, Alvin Chang, *The Facebook and Cambridge Analytica Scandal, Explained with a Simple Diagram*, VOX (May 2, 2018, 3:25 PM), https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram [https://perma.cc/7S42-T2QH].

154. *See* Steven Melendez, *How Google is Breaking EU Privacy Law, According to a New Complaint*, FAST COMPANY (Sept. 13, 2018), https://www.fastcompany.com/90236273/google-faces-gdpr-privacy-complai nt-over-its-targeted-ads-from-brave-browser [https://perma.cc/6JU3-2YGK].

155. Complaint, Snapchat, Inc., 79 Fed. Reg. 27611 (F.T.C. May 8, 2014) (FTC File No. 132 3078), https://www.ftc.gov/system/files/documents/cases/140508snapchatcmpt.pdf [https://perma.cc/5CFF-YWC9] [hereinafter Snapchat Complaint].

156. *See* Kate Darling, *Extending Legal Protection to Social Robots*, *in* ROBOT LAW 214 (Ryan Calo, A. Michael Froomkin & Ian Kerr eds., 2016) (defining social robots as physically embodied agents that interact with and learn from humans on a social level).

157. *See PARO Therapeutic Robot*, PARO, http://www.parorobots.com/ [https://perma.cc/KU5K-USZH].

158. *See* AIBO, https://us.aibo.com/ [https://perma.cc/KKN8-VMJL].

159. Cynthia Breazeal, *Toward Sociable Robots*, 42 ROBOTICS & AUTONOMOUS SYS. 167, 168 (2003).

160. *See* M. Ryan Calo, *People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship*, 114 PENN ST. L. REV. 809, 817–25 (2010) [hereinafter Calo, *People Can Be So Fake*] (discussing how cyberlaw has traditionally focused on data collection as the salient privacy problem posed by robotics); Woodrow Hartzog, *Unfair and Deceptive Robots*, 74 MD. L. REV. 785, 797–801

surveillance,[161] suffocate autonomy,[162] and eliminate opportunities for solitude.[163] They can spy on us and nudge us to buy things.[164] And our connections to social robots can be leveraged to manipulate us into paying for high-priced upgrades or responding to advertisements snuck into answers and responses.[165]

As Ryan Calo has argued, social robots have three essential qualities that help us understand the risks they pose to our privacy: embodiment, emergence, and social valence. Social robots are embodied in that they occupy physical form. They may be programmed to act based on a coded series of ones and zeros, but we phenomenologically experience social robots taking physical action in the physical world.[166] Robots are also emergent in that they (and their programming) can learn and adapt to new circumstances and new demands.[167] To us, then, social robots are part of an ongoing social dance of back-and-forth interaction much like humans.[168] And they have a social valence in that and we tend to use social models to understand them.[169] This is why we tend to become uncomfortable with and resistant to behaviors that would "harm" robots;[170] their shape, verbal skills,

(2015) [hereinafter Hartzog, *Unfair and Deceptive Robots*] (discussing all the ways robots can spy on their human users).

161. *See* Freed et al., *supra* note 7.

162. *See* Calo, *People Can Be So Fake*, *supra* note 160, at 847; *see also* M. Ryan Calo, *Robots and Privacy*, *in* ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTS 187, 195 (Patrick Lin et al. eds., 2012); M. Ryan Calo, *The Drone As Privacy Catalyst*, 64 STAN. L. REV. ONLINE 29 (2011) (backlash around robots bound up with cultural depictions of robots); M. Ryan Calo, *Against Notice Skepticism in Privacy (and Elsewhere)*, 87 NOTRE DAME L. REV. 1027 (2012) [hereinafter Calo, *Against Notice Skepticism*] (robotic surveillance introduces cues of surveillance that are actually missing in computer-mediated cyberspace).

163. *See* Calo, *People Can Be So Fake*, *supra* note 160, at 843–46.

164. *See* Hartzog, *Unfair and Deceptive Robots*, *supra* note 160, at 797–802.

165. *See* Darling, *supra* note 156, at 221.

166. *See* Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 532–37 (2015) [hereinafter Calo, *Cyberlaw*]. Phenomenologists argue that the only way to understand the world around us is to start from the point of human embodied perception and experience, rather than from an ontological perspective. *See* MAURICE MERLEAU-PONTY, PHENOMENOLOGY OF PERCEPTION xi–xii (Ted Honderich ed., Colin Smith trans., 1962) ("The world is not an object such that I have in my possession the law of its making; it is the natural setting of, and field for, all my thoughts and all my explicit perceptions."). Mark Coeckelbergh has taken a phenomenological approach to robots, arguing that ethics for robotics should start at the point of human perception rather than technical definitions of robots. *See, e.g.*, Mark Coeckelbergh, *Robot Rights? Towards a Social-Relational Justification for Moral Consideration*, 12 ETHICS & INFO TECH. 209 (2010); Mark Coeckelbergh, *Humans, Animals, and Robots: A Phenomenological Approach to Human-Robot Relations*, 3 INT'L J. OF SOC. ROBOTICS 197 (2010); Mark Coeckelbergh, *Virtual Moral Agency, Virtual Moral Responsibility*, 24 AI & SOCIETY 181 (2009).

167. *See* Calo, *Cyberlaw*, *supra* note 166, at 539.

168. *See* GOFFMAN, EVERYDAY LIFE, *supra* note 11, at 62–63.

169. *See* Calo, *Cyberlaw*, *supra* note 166, at 545–46.

170. *See, e.g.*, Peter H. Kahn, Jr. et al., *The New Ontological Category Hypothesis in Human-Robot Interaction*, 2001 6TH INT'L CONF. ON HUMAN-ROBOT INTERACTIONS 159 (2001).

adaptability, and seemingly autonomous actions distinguish them from machines, appliances, and tools in our minds.[171]

For these reasons, social robots create social spaces that put our privacy at risk. We interact with social robots in ways similar to interacting with humans because regardless of our intellectual ability to recognize that robots aren't human, robots with human qualities *feel* social to us. This isn't accidental. Social robots are specifically designed to trigger our predisposition to anthropomorphize,[172] and to induce the kind of trust we usually reserve for other humans. This lulls us into a false sense of security, thus increasing our propensity to disclose personal information to a data-hungry corporation hiding behind a social veil. Woodrow Hartzog has called this "the most fundamental reason we are vulnerable to robots."[173]

These problems exist in part because online social networks and social robots are built to encourage disclosure. But instead of supporting disclosure and user safety with trust-building design, technosocial spaces leverage design to cue false trust among their members, putting up a veneer in front of massive, invasive data collection and sharing with third parties. And the law lets them do it. Therefore, online social networks and social robots create unsafe social spaces.

B. Organic Trust in Technosocial Spaces

All of the social forces that naturally generate trust among individuals offline are present in technosocial spaces. Our propensity to disclose information on online social networks is positively correlated with the number of friends—and even more so with the number of close friends—we have on the platform.[174] In one study, eighty-five percent of users said would accept a Facebook "friend" request from a stranger if they shared a sufficient number of mutual friends and eighty-one percent would do the same if the stranger was friends with their close friends.[175] The vast majority of users who identify as LGBTQ would accept a friend request from a stranger simply because they also identified as queer[176] and gay and bisexual men tend to trust strangers on queer-only dating apps because of the

171. Don Ihde has argued that robots' meaning in society is "multistable": we may sometimes see robots as machines and sometimes see them as more than machines. *See* DON IHDE, TECHNOLOGY AND THE LIFEWORLD (1990).

172. Darling, *supra* note 156, at 214.

173. Hartzog, *Unfair and Deceptive Robots*, *supra* note 160, at 791.

174. *See* Waldman, *Privacy, Sharing, and Trust*, *supra* note 61, at 216–17.

175. *Id.* at 219.

176. *Id.* at 219–220.

platform's inherent queerness,[177] thus lending credibility to the theory that sharing an out-group identity can build trust.

Users also build trust through experience, over time, and via interaction. Dating app users tend to share images of themselves after "chatting with the other person" for a time, ranging from a few hours to a few weeks. Reciprocity also inspires trust and, thereby, sharing of personal information. Queer users of online social networks believe strongly in reciprocal sharing because the mutual surveillance it allows can inspire trust and thus mitigate the risks inherent in sharing nude or seminude images.[178]

Trust generation is even more organic in social spaces involving social robots. We have an innate, evolutionary need to connect with others,[179] and we naturally apply social models to understand and interact with the world around us.[180] One of those social models is anthropomorphization, or ascribing human characteristics to nonhuman things, like when we talk to our dogs or stuffed animals or when Tom Hanks paints a face on a volleyball in *Cast Away*.[181]

Our natural, "inborn" tendency to connect with social robots[182] grows stronger as objects take on more humanlike characteristics. Hanks painted a face on his volleyball, which only then could be perceived as a head. Many people put eyes on their Roombas,[183] and two-thirds of Roomba owners give them names.[184] Human users bonded with ELIZA, a computer psychoanalysis program that asked users questions like a therapist and filled in conversations with dummy placeholder comments, prompting the lead experimenter to issue a warning call about the manipulative effects of artificial intelligence.[185] And another study showed that humans will engage happily and politely with kind and polite computer programs.[186]

Since social robots have more human*ish* qualities than Roombas or desktop computers, the trust that develops between humans and social robots is likely even more powerful than the trust developed in these

177. *See* Waldman, *Law, Privacy, and Online Dating*, *supra* note 6.

178. *Id.*

179. There is a long literature on this going back decades. *See, e.g.*, Roy F. Baumeister & Mark R. Leary, *The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation*, 117 PSYCHOL. BULL. 497 (1995); Abraham H. Maslow, *A Theory of Human Motivation*, 50 PSYCHOL. REV. 370 (1943).

180. *See* MERLEAU-PONTY, *supra* note 166.

181. CAST AWAY (20th Century Fox 2000).

182. *See* Calo, *People Can Be So Fake*, *supra* note 160, at 811, 826.

183. This practice has reached popular culture. *See American Dad!: May the Best Stan Win* (Fox television broadcast Feb. 14, 2010).

184. *See* Robert Boyd, *Robots are Narrowing the Gap With Humans*, MCCLATCHY NEWSPAPERS (Apr. 27, 2009), https://www.mcclatchydc.com/news/politics-government/article24534961.html [https://perma.cc/PN3K-4G64].

185. Calo, *People Can Be So Fake*, *supra* note 160, at 836.

186. *Id.* at 837.

experiments. Sherry Turkle has shown that people of all ages establish tight bonds with social robots. After playing with a human*ish* robot that could make eye contact, follow a child around, and imitate the child's movements, an eleven-year-old called her social robot "something that's part of you, . . . something you love, kind of like another person, like a baby."[187] A seventy-one-year-old user of a furry robot that looked like a koala bear also noted that "[w]hen I looked into his large, brown eyes, I fell in love after years of being quite lonely . . . [and] I swore to protect and care for the little animal."[188] Other studies have shown that people playing prisoner dilemma games with technological interfaces tended to keep their promises with more human*ish* partners.[189] There are a growing number of persuasive studies just like these.[190] Suffice it to say, as Karl MacDorman and Hiroshi Ishiguro note, "[h]uman-like appearance and behavior . . . elicit the sorts of responses that people typically direct toward one another."[191] And the more human*ish* they get, "the more human-directed . . . expectations are elicited."[192] Central to those expectations is trust.

C. Manipulative Designs that Entice Disclosure

Rather than using design to counterbalance powerful norms of disclosure to support organic trust, technosocial spaces leverage design to elicit more disclosure by creating a veneer of false trust where actually none exists. There are many examples of this,[193] but the designs of Snapchat, Facebook News Feeds, and embodied social robots are paradigmatic.

Snapchat was originally designed to manipulate disclosure by creating the appearance of trust. It sold itself as a privacy-protective platform by highlighting a design choice that made any image or video, or "snap," sent

187. Sherry Turkle, *In Good Company? On the Threshold of Robotic Companions, in* CLOSE ENGAGEMENTS WITH ARTIFICIAL COMPANIONS: KEY SOCIAL, PSYCHOLOGICAL, ETHICAL, AND DESIGN ISSUES 3 (Yorick Wilks ed., 2010).

188. *Id.* at 5.

189. Salvatore Parise et al., *Cooperating with Life-Like Interface Agents*, 15 COMPUTS. IN HUM. BEHAV. 123, 124 (1999).

190. For an in-depth discussion of many of the field experiments on human-computer and human-robot interaction, see Calo, *People Can Be So Fake*, *supra* note 160, at 840.

191. Karl F. MacDorman & Hiroshi Ishiguro, *The Uncanny Advantage of Using Androids in Cognitive and Social Science Research*, 7 INTERACTION STUD. 297, 316 (2006).

192. *Id.* at 309.

193. Woodrow Hartzog collects and reports on a plethora of examples in HARTZOG, PRIVACY'S BLUEPRINT, *supra* note 21, at 197–275. There is also a growing research agenda on the impact of "dark patterns" on disclosure online. Dark patterns are "user interface design choices that benefit an online service by coercing, steering, or deceiving users into making unintended and potentially harmful decisions." Arunesh Mathur, *Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites*, ACM CSCW, 2019, at 1, https://arxiv.org/pdf/1907.07032.pdf [https://perma.cc/U93F-S37P]; *see also* Ryan Calo, *Digital Market Manipulation*, 82 GEO. WASH. L. REV. 995 (2014) (arguing, in relevant part, that consumer protection law must adapt to address the market manipulations of dark patterns).

across its platform would automatically disappear after several seconds and could not be retained by users. Except, it was not originally designed that way. Before sending a snap, users were shown a screen that required them to designate the amount of time the snap will survive before disappearing, thus building expectations of confidentiality and discretion and a sense of user control.[194] Snaps could not be sent without selecting an option. In reality, there were several ways snaps sent could be saved, downloaded, or copied.[195] This gave users the false impression, reinforced in the platform's product descriptions and Frequently Asked Questions,[196] that they actually had control over what their recipients could do with their snaps.

There was no trust designed into Snapchat. Until October 2013, it stored all videos in unprotected spaces on users' phones, which allowed recipients to simply search for and download a video they wanted to save.[197] Snapchat also allowed any third-party application to access its application programming interface and download or copy videos and images.[198] Not only were these vulnerabilities not conveyed to users, but the platform's design created contrary expectations.

Facebook designs its interface to deliver cues of trust to its members so they will share more personal information: it creates a sense of community through rich profiles, privileges our friends' posts so we see them first, and publicly informs us of our friends' online behavior to encourage reciprocal sharing.[199] Facebook's News Feed, the running list of stories and posts from our friends, is designed to make it difficult for users to distinguish between social posts and native advertisements. Among other design tactics, both types of posts are prefaced by notices about our friends' interactions—"Jane, Joe, and 18 others liked this"—and both are followed by notifications of our friends' comments—"David, Maggy, and 27 others commented on this post." This design cues trust: users can look to Jane, Joe, David, and Maggy and feel confident that the post is social, meaningful, and relevant. But when the same trust cues appear on an advertisement, on a link to a third

194. Snapchat Complaint, *supra* note 155, at ¶6.

195. *Id.* at ¶¶ 9–17. Much of the FTC's case against Snapchat focused on the company's failure to disclose certain data collection practices in its privacy statement. *See id.* ¶¶ 8–33. But broken promises litigation is just one part of the FTC's privacy jurisprudence. *See* Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 667 (2014). As Solove & Hartzog point out, the FTC has developed a broader view of unfair or deceptive practices, including, for example, "deception by omission," *id.* at 631, "inducement" to share personal information, *id.* at 632–33, and "pretexting," *id.* at 633, to name just a few. Their persuasive argument is that "through a common law-like process, the FTC's actions have developed into a rich jurisprudence that is effectively the law of the land for businesses that deal in personal information." *Id.* at 589.

196. Snapchat Complaint, *supra* note 155, at ¶¶ 7–8.

197. *Id.* at ¶ 10.

198. *Id.* at ¶ 11.

199. *See* Waldman, *Privacy, Sharing, and Trust*, *supra* note 61, at 221–23.

party whose data collection practices are unknown, or, worse yet, on click bait to a radically invasive quiz or website,[200] the design transfers the trust we have in our friends to a third-party advertiser about which we know little. These design tactics hide privacy risks from users, cuing trust where no protections for users actually exist.

Social robots also use design to cue false trust. By creating machines that mimic human behavior while looking like adorable baby seals or having other anthropomorphic characteristics, social robots lull us into a false sense of trust. And they tend to manipulate the most vulnerable among us: the elderly, the disabled, and the lonely.[201] Advanced social robots use sounds, language, and movement in ways reminiscent of, if not identical to, humans. As such, social robots have the capacity to appear trustworthy; they can be partners in ongoing social interaction. For Erving Goffman, social interaction is an ongoing dance, with both leaders and followers sharing responsibility for continuing the dance: individuals determine what they want to reveal and their partners (Goffman calls them the "audience") keep up the charade of persona management by playing along.[202] Social robots play along, as well. They generally behave in expected ways: Alexa answers questions,[203] Roombas skate along the floor, robot butlers get coffee. We are meant to experience them phenomenologically and, therefore, see the designed-in human*ish* qualities rather than the technology company lurking behind the curtain.

They are, then, classic "Wizard of Oz" setups with unique abilities to harm.[204] As Jacqueline Kory Westlund and Cynthia Breazeal note, users "may not realiz[e] that a human is hearing everything they say."[205] "Given that social robots are designed to draw us in, often engaging us emotionally and building relationships with us, the robot itself could be deceptive in that

200. *See, e.g.*, *Think Before You Click: How Facebook Clickbait Puts Users at Risk*, WRAL.COM (May 10, 2016), http://www.wral.com/think-before-you-click-how-facebook-clickbait-puts-users-at-risk-/15682285/ [https://perma.cc/4SSC-QJAL]; Claire Suddath, *The Weather Channel's Secret: Less Weather, More Clickbait*, BLOOMBERG (Oct. 9, 2014, 5:00 AM), https://www.bloomberg.com/news/articles/2014-10-09/weather-channels-web-mobile-growth-leads-to-advertising-insights [https://perma.cc/B9KU-DVTU].

201. *See* James A. Mourey et al., *Products as Pals: Engaging with Anthropomorphic Products Mitigates the Effects of Social Exclusion*, 44 J. CONSUMER RES. 414 (2017) (finding that social robots tend to elicit the greatest response from people who are lonely or recently suffered loss).

202. GOFFMAN, EVERYDAY LIFE, *supra* note 11, at 9.

203. Alexa does not normally laugh fiendishly at 10:00 P.M., and the fact that it behaved out of the ordinary is what struck many users as problematic when it did laugh. *See* Brian Koerber, *Amazon Reveals Why Alexa is Randomly Laughing and Creeping People Out*, MASHABLE (Mar. 7, 2018), https://mashable.com/2018/03/07/wh y-amazon-alexa-laughing/ [https://perma.cc/Z28J-7U2U].

204. *See* Hartzog, *Unfair and Deceptive Robots*, *supra* note 160, at 804–05.

205. Jacqueline Kory Westlund & Cynthia Breazeal, *Deception, Secrets, Children, and Robots: What's Acceptable?*, 10TH ACM/IEEE CONFERENCE ON HUMAN-ROBOT INTERACTION, 2015, at 1, www.openroboethics.org/hri15/wp-content/uploads/2015/02/Mf-Westlund.pdf [https://perma.cc/KTR7-LU56].

it appears to" make a connection, "but 'in reality,'" it's just gathering data on us.[206] Ian Kerr presciently recognized this over a decade ago: "Like Hollywood's finest directors, who are able to steer their audiences' attention away from the false assumptions that they have so skillfully engendered, some software programmers are applying principles of cognitive science to develop electronic entities that garner consumer trust. Unfortunately, some e-businesses are exploiting these applications to garner trust where no such trust is warranted."[207]

D. Legal and Regulatory Void

Technology companies that design technosocial platforms are allowed to leverage design in manipulative ways because, for the most part, the law allows it. Legal regimes like malpractice law, tort law, trademark law, and private ordering schemes like professional licensing requirements constrain the opportunistic behavior of information trustees in the offline context.[208] Digital spaces do not just lack comparable legal levers. The laws that do exist actively promote insecurity and risk to users.

Online social networks have no legal incentive to police their platforms for hate and harassment because Section 230 of the Communications Decency Act immunizes them for most third party conduct on their platforms.[209] Section 230 was passed in reaction to a lawsuit against Prodigy, one of the original online service providers.[210] At the time, Prodigy filtered profanity and other harmful content out of its platform, holding itself out as a family-friendly technosocial space.[211] After a user posted defamatory comments about a securities company, the firm sued Prodigy, arguing that Prodigy should be liable as publisher of the defamation.[212] The New York Supreme Court agreed.[213]

Section 230 was then introduced in Congress by then-Representative Christopher Cox of California and now-Senator Ron Wyden of Oregon to overturn the Prodigy case and to protect internet service providers from lawsuits focusing on imperfect filtering.[214] Congress stated that it passed the

206. *Id.*

207. Ian R. Kerr, *Bots, Babes, and the Californication of Commerce*, 1 U. OTTAWA L. & TECH. J. 285, 288 (2004).

208. *See supra* Part II.

209. *See* 47 U.S.C. § 230(c)(1).

210. *See* Stratton Oakmont, Inc. v. Prodigy Servs., Co., No. 31063/94, 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995).

211. *Id.* at *2.

212. *Id.* at *1–2.

213. *See* Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401, 405 (2017).

214. *See* 141 CONG. REC. H8460-01 (Aug. 4, 1995).

law to preserve the internet as "a forum for a true diversity" of views "with a minimum of government regulation," and to maintain "the vibrant and competitive free market that presently exists for the Internet . . . unfettered by Federal or State regulation."[215]

Congress's other reason for enacting Section 230 was to encourage Internet intermediaries, users, and parents to self-police the Internet for obscene conduct.[216] But in interpreting the clause, federal courts cemented broad immunity for platforms, leaving no legal incentive to police bad behavior. In *Zeran v. America Online*,[217] for example, the Fourth Circuit noted that lawsuits against providers for third-party content would risk "freedom of speech in the new and burgeoning Internet medium."[218] In a an expansive, maximalist holding, the court stated that "Section 230 was enacted, in part, to maintain the robust nature of Internet communication, and accordingly, to keep government interference in the medium to a minimum."[219] This and other broad immunity decisions created an online world where platforms have no legal incentive to make their platforms safe and where perpetrators do not fear the consequences of their actions.

Nor does current law encourage online social networks or the designers of social robots to protect user privacy. Although new privacy laws like Europe's General Data Protection Regulation (GDPR)[220] and California's Consumer Privacy Act,[221] and the renewed prospect for comprehensive privacy legislation in the United States,[222] offer the promise of reasonable constraints on predatory data collectors, United States consumers are still operating under a notice-and-choice regime that carries with it minimal obligations for technology companies.[223] Notice-and-choice requires

215. 47 U.S.C. §§ 230(a)(3)–(4), (b)(2).

216. *See* § 230(b)(4); 141 CONG. REC. H8469-70 (Aug. 4, 1995) (statements of Reps. Cox, Wyden and Barton); Blumenthal v. Drudge, 992 F. Supp. 44, 52 (D.D.C. 1998). This is sometimes called the "Good Samaritan" clause. Kate Klonick argues that social media platforms have economic incentives to take this to heart when it comes to harassing, hateful, and harmful speech. *See* Klonick, *supra* note 10, at 1627–30.

217. 129 F.3d 327 (4th Cir. 1997).

218. *Id.* at 330.

219. *Id.* at 330; *see also* Ben Ezra, Weinstein, & Co. v. Am. Online, Inc., 206 F.3d 980, 985 n.3 (10th Cir. 2000) (similar).

220. *See* Regulation 2016/679, 2016 O.J. (L 119) [hereinafter GDPR].

221. *See* Assem. B. 375, 2017–2018 Reg. Sess. (Cal. 2018).

222. Senator Brian Schatz (D-HI) has introduced a draft bill of federal privacy legislation. Data Care Act, S. 3744, 115th Cong. (2018), https://www.congress.gov/bill/115th-congress/senate-bill/3744 [https://perma.cc/4PWJ-B4RV]. Technology companies are hoping to pre-empt the CCPA with a federal law written with their input. *See* Cecilia Kang, *Tech Industry Pursues a Federal Privacy Law – On its Own Terms*, N.Y. TIMES (Aug. 26, 2018), https://www.nytim es.com/2018/08/26/technology/tech-industry-fed eral-privacy-law.html [https://perma.cc/65RV-B22G].

223. That is, unless those companies fall within one of the few industries regulated by sector-specific privacy laws, such as the Health Insurance Portability and Accountability Act (HIPAA) or Gramm-Leach-Bliley, which regulates financial information. *See* Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (1996) (codified as amended at 42

companies that collect our data to describe their data collection tactics and data use practices, and to give us the opportunity to opt out.[224] Privacy policies are, therefore, its primary legal tool. As a doctrine of informed consent,[225] notice-and-choice is supposed to give us control over our data by giving us the information we need to make rational disclosure decisions. But the regime is inadequate. Not only are privacy policies incomprehensible,[226] but the entire endeavor of notice-and-choice is premised on the myth of an autonomous user.[227] We do not make perfectly rational disclosure decisions, and we wouldn't even if we could comprehend privacy policies.[228] Far from providing any real notice, the consent model employed today actually disempowers users by putting them in a position in which their consent—"Click 'I agree' to continue"—is used as an excuse to allow companies to use data how they see fit. Consent, then, extracts power that should belong to users.[229]

And the prospect of change under the GDPR is slim. Some scholars have argued that the GDPR is little more than the old notice-and-choice regime

U.S.C. §§ 1320d(1)–(9)); 45 C.F.R. § 164.528 (2016); Gramm–Leach–Bliley Act (GLBA), Financial Services Modernization Act of 1999, Pub. L. 106–102, 113 Stat. 1338 (1999) (codified as amended at 15 U.S.C. §§ 6801–6809). But even those laws are incomplete. HIPAA only protects certain health data held by certain covered entities, like health insurance plans, clearinghouses, HMOs, and company health plans. And it only applies to doctors if they electronically transfer information in connection with a transaction for which the Department of Health and Human Services has adopted a standard. *See* 45 C.F.R. §§ 160.102–160.103; *see also Covered Entities and Business Associates*, HHS.GOV, http://www.hhs.gov/hipaa/for-professionals/covered-entities/ [https://perma.cc/2UG7-NWPB]. And the GLBA only applies to companies that offer consumer financial products or services to explain their information-sharing practices. It does not cover the entire financial services industry. *See* 113 Stat. at 1443–45 (codified at 15 U.S.C. § 6809 (2000)).

224. *See* Solove & Hartzog, *supra* note 195, at 592.

225. *See* Joel R. Reidenberg et al., *Privacy Harms and the Effectiveness of the Notice and Choice Framework*, 11 I/S: J. L. & POL'Y FOR INFO. SOC'Y 485, 518 (2015). The principle of informed consent, as in the analogous contexts of medical procedures and scientific research, flows directly from Kant's categorical imperative: "Act in such a way as to treat humanity, whether in your own person or in that of anyone else, always as an end and never merely as a means." IMMANUEL KANT, GROUNDWORK FOR THE METAPHYSIC OF MORALS 29 (2005); *see also* Jorge L. Contreras, *Genetic Property*, 105 GEO. L.J. 1, 18 (2016).

226. *See* Joel R. Reidenberg et al., *Disagreeable Privacy Policies: Mismatches Between Meaning and Users' Understanding*, 30 BERKELEY TECH. L.J. 39, 40, 87–88 (2015) (presenting results of an experimental study showing that average internet users do not understand privacy policies and that even experts cannot agree on the meanings of certain terms).

227. *See, e.g.*, JULIE E. COHEN, CONFIGURING THE NETWORKED SELF: LAW, CODE, AND THE PLAY OF EVERYDAY PRACTICE 16–21 (2012) (as part of the governing principles of cyberspace); Cohen, *DRM and Privacy*, *supra* note 19, at 225–27 (users are constrained by the built online environments around them); MICHAEL J. SANDEL, DEMOCRACY'S DISCONTENT: AMERICA IN SEARCH OF A PUBLIC PHILOSOPHY 25–28 (1996) (as the foundation of political philosophy).

228. Acquisti, John & Loewenstein, *supra* note 61, at 160.

229. *See* HARTZOG, PRIVACY'S BLUEPRINT, *supra* note 21, at 207–211 (discussing the problem of "extracted consent").

with a few new bells and whistles.[230] Some of its language is impossibly vague and of little help in supporting and fostering safe sharing online.[231]

Social robots exist in a similar regulatory void. Their data use practices, at least in the United States, are governed by a notice-and-choice regime, and they don't even have webpage interfaces on which one can find a privacy policy. And, although the FTC has the power to regulate the manipulative design strategies of social robots,[232] it has yet to exercise that authority in any significant way.[233] That could be because social robots are so new. It could also be because the FTC continues to defer to a self-regulation privacy regime, which gives technology companies the power to set the terms of industry practice.

IV. PROPOSED CHANGES TO DESIGN AND LAW

So far, I have argued that law and design must work together to buttress trust norms that protect socially beneficial disclosures. I have also shown that current law and design thinking do the opposite when it comes to technosocial spaces. They are leveraged to promote disclosures that benefit technology companies, but put users at risk. We can change that. We need to design our technosocial spaces to build trust endogenously, and we need law to both incent technology companies to design for privacy and safety and protect users when something goes wrong.

A. Designing for Trust and Safety

Technosocial spaces can be designed to enhance trust, protect our privacy, and keep us safe. As noted earlier, "design" refers to the set of processes throughout which new technologies are built and come about.[234] As such, both technical specifications and corporate practices matter.

230. *See* Woodrow Hartzog, *The Inadequate, Invaluable Fair Information Practices*, 76 MD. L. REV. 952, 955–56 (2017) (calling the GDPR a "FIPs-based law[]").

231. *See, e.g.*, Margot E. Kaminski, *The Right to Explanation, Explained*, 34 BERKELEY TECH. L.J. 189, 195 (2019) (arguing that GDPR lays out several "broad standards" that need to be given "specific substance over time"); Waldman, *Privacy's Law of Design*, *supra* note 26, at 1255–56 (examining the ambiguities in GDPR Article 25); Alison Cool, *Europe's Data Protection Law Is a Big, Confusing Mess*, N.Y. TIMES (May 15, 2018), https://www.nytimes.com/2018/05/15/opinion/gdpr-europe-data-protection.html [https://perma.cc/DNM3-XLNN] (calling the GDPR "staggeringly complex" and "intentionally ambiguous").

232. *See* 15 U.S.C. § 45(a)(1) ("Unfair methods of competition in or affecting commerce, and unfair or deceptive acts or practices in or affecting commerce, are hereby declared unlawful."). The FTC was given the authority to prevent such practices in subsection (a)(2). *See* 15 U.S.C. § 45(a)(2).

233. *But see* Hartzog, *Unfair and Deceptive Robots*, *supra* note 160, at 812–21 (arguing that the FTC both has the power and the capacity to address deceptive robots); *see also* HARTZOG, PRIVACY'S BLUEPRINT, *supra* note 21, at 131–34 (calling on the FTC to take more aggressive steps to regulate deceptive design, generally).

234. *See* HARTZOG, PRIVACY'S BLUEPRINT, *supra* note 21, at 11.

On the technology side, online social networks can change defaults to better protect our privacy. On Venmo, a social network for online transactions, interactions are public by default "because it's fun" to see what our friends are buying, the company says.[235] Corporate perceptions of user entertainment aside, this choice only makes sense in a world where privacy gets short shrift in design. Settings should default to privacy-protective options, giving users the power to determine for themselves what information they will share, when they share it, and with whom. This is one of the principles behind privacy by design.

Although I am skeptical that a notice-and-choice regime could ever do much good, notice can be improved. Warnings about privacy risks could reach users on a visceral or emotional level rather than merely through incomprehensible privacy policies.[236] This would be especially useful in technosocial spaces involving social robots, where websites are absent. On social networks, sponsored posts on Facebook could use distinguishing colors and presentation to separate them from social posts from our friends.[237] So-called "just in time" notifications, or popups that inform users of data collection as it happens, could help users make better, in-the-moment decisions.[238] And social websites could be redesigned to require users to opt-in to data collection, as required in some circumstances by the GDPR.[239] Remaining opt-out tools that protect users from behavioral tracking and cookies could be built in to browsers, obviating the need to opt-out of data collection every time we visit a website.[240]

Safety by design could involve ephemeral messaging for intimate images, access restrictions, streamlined takedown procedures, and frictionless tagging of profiles that spew hate, engage in harassment, and violate other terms of use.[241] For example, the queer-oriented dating app Scruff allows users to easily tag profiles that are racist, hateful, and

235. Marrian Zhou, *Venmo Transactions Are Public By Default as Part of It's Social Strategy*, CNET (July 20, 2018, 2:22 PM), https://www.cnet.com/news/venmo-explains-why-transactions-are-public-by-default/ [https: //perma.cc/79N9-NG5W].

236. *See* Calo, *Against Notice Skepticism*, *supra* note 162, at 1030 (calling for the use of "visceral" notice to enhance user understanding of privacy risks).

237. *See* Waldman, *Privacy, Sharing, and Trust*, *supra* note 61, at 226.

238. *Id.* at 226–27.

239. *See, e.g.*, *Guidelines on Consent Under Regulation 2016/679*, at 12, 16 (Nov. 28, 2017), https://iapp.org/media/pdf/resource_center/20180416_Article29WPGuidelinesonConsent_publishpdf.pdf [https: //perma.cc/T9QP-ZXHY].

240. *See Commission Proposal for a Regulation 2017/0003 of the European Parliament and of the Council Concerning the Respect for Private Life and the Protection of Personal Data in Electronic Communications and Repealing Directive 2002/58/EC*, COM (2017) 10 final (Oct. 1, 2017), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017PC0010&from=EN [https://perma.cc/YZ5R-JQBS].

241. *See* HARTZOG, PRIVACY'S BLUEPRINT, *supra* note 21, at 228–29; *see also* Waldman, *Law, Privacy, and Online Dating*, *supra* note 6.

harassing. Reporting procedures are in-app, meaning users do not have to click through to a second or third webpage to find a reporting form. The information necessary is also limited: the flag and a brief explanation of the incident or pattern of behavior are sufficient to initiate Scruff's forty-eight-hour window for responding to all flags and harassment claims. That frictionless reporting regime may be at least part of the reason why Scruff has few, if any, problems with racist profiles and revenge pornography.[242]

Designing for trust also requires broad corporate commitments that reach into the daily work of privacy and safety. Elsewhere, I have argued that robust visions of privacy may not make their way into technology product design because some designers—the engineers on the ground—do not share a commitment to privacy.[243] Nor do they have the educational and conceptual tools to spot ambiguous privacy issues raised by code.[244] Technology companies can address part of this problem endogenously by integrating privacy and law into the design process. Companies should include lawyers, sociologists, and other social scientists in design meetings. At a minimum, these experts could be co-located with engineering teams. And privacy should be included in corporate missions and designer training, reinforced throughout their term of employment. In short, privacy has to be part of both the ethos and routine of designers and their employers.[245]

Social platforms can also reorient their content moderation policies toward protecting marginalized populations, sexual privacy,[246] and the safety of their users. Content moderation is a mixture of technology design (artificial intelligence) and human capital resources[247] deployed to create social environments for their users.[248] Kate Klonick has argued that platforms try to moderate content to match the expectations of their users, which include protections for free speech as well as online safety.[249] But much of that work is influenced by traditional First Amendment doctrine: the people who conceptualized social media's place in the speech governance ecosystem and developed moderation policies were lawyers steeped in First Amendment law they had learned in school and practiced in

242. *See* Waldman, *Law, Privacy, and Online Dating*, *supra* note 6, at 16–17.

243. *See* Ari Ezra Waldman, *Designing Without Privacy*, 55 HOUSTON L. REV. 659, 662–63 (2018).

244. *See id.* at 716–25.

245. *See* Martha S. Feldman & Brian T. Pentland, *Reconceptualizing Organizational Routines as a Source of Flexibility and Change*, 48 ADMIN SCI. Q. 94, 94–95 (2003) (identifying that corporate executives are responsible for the "ostensive" aspect of work, including setting the mission, and workers on the ground "perform" the routine of work, translating that mission into actions and products); *see also* Bruno Latour, *The Powers of Association*, 32 SOC. REV. 264, 266–68, 271–73 (1984).

246. *See* Citron, *Sexual Privacy*, *supra* note 18.

247. *See* Klonick, *supra* note 10, at 1635–47.

248. *Id.* at 1602.

249. *Id.* at 1627–28.

the courtroom.[250] It should come as no surprise that content moderation started from a "baseline" of liberal free speech norms.[251] Moderation policies could instead be designed to tip the scales in favor of safety rather than unfettered speech. Because the underlying policies still remain the same, recent steps to ban right-wing hate are welcome, yet woefully insufficient steps in this regard.[252]

B. Law to Support Trust and Safety

Technology companies need legal incentives to design for real trust and safety, especially where such design may complicate data collection in an economy dependent upon it.[253] That may be a heavy lift, but reasonable changes are necessary to protect users. This section describes four legal levers we can push: fiduciary law, FTC enforcement, privacy by design, and reform to Section 230.

1. Information Fiduciaries

If we want technosocial spaces to inspire trust and foster safe and socially beneficial disclosure in ways similar to AA meetings, teams of coworkers, and attorney-client relationships, then we should treat technology platforms as information fiduciaries.[254] Fiduciaries have special obligations of loyalty because we put our trust in them. Estate managers, investment advisers, lawyers, and doctors are classic examples of fiduciaries: they handle our money, secrets, and lives under duties of loyalty and care.[255] As Jack Balkin has observed, fiduciary duties are "duties of trust;" the word *fiduciary* even comes from the Latin word for *trust*.[256] And, as we have seen, we share information with others in contexts of trust.[257] Even Justice Gorsuch recognized this point in his dissent in *Carpenter v. United States*,[258] where he challenged the Third Party Doctrine's erosion of privacy by noting that "[p]eople often do reasonably expect that information they entrust to third

250. *Id.* at 1625–29.
251. *Id.* at 1618.
252. *See* April Glaser, *Why Facebook's Latest Ban Was So Underwhelming*, SLATE (May 3, 2019, 7:39 PM), https://slate.com/technology/2019/05/facebook-alex-jones-ban-underwhelming.html [https://perma.cc/ZUQ4-CBEZ].
253. *See* HARTZOG, PRIVACY'S BLUEPRINT, *supra* note 21, at 197.
254. *See* Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183 (2016).
255. *Id.* at 1207–08.
256. *Id.* at 1208.
257. *See also* WALDMAN, PRIVACY AS TRUST, *supra* note 24, at 49–60.
258. 138 S. Ct. 2206 (2018).

parties . . . will be kept private."[259] A bill proposed at the end of the 115th Congress by Senator Brian Schatz of Hawaii reflected some of these ideas.[260]

Technology companies running technosocial platforms should be considered information fiduciaries for the same reasons that doctors, lawyers, and investment advisers are considered traditional fiduciaries. We are vulnerable to them because they know everything about us. We are dependent on them because of the services they provide and the expertise they bring to those services. And they hold themselves out as sufficiently trustworthy to gain our business.[261]

This would have profound effects on users, design, and corporate policies. Information fiduciaries should, first and foremost, never act like "con men." Conning users would be like using dark patterns to purposely extract personal information against user wishes or Google Maps holding itself out as providing the best or fastest route from JFK International Airport to the West Village and then delivering a route that drives passed a Chipotle because Chipotle paid Google $100. Information fiduciaries would not be able to leverage data in ways that violate social norms. Nor could they use our data to manipulate, lie, or take away our freedom.[262] An information fiduciaries approach would let us use our credit cards knowing our purchasing histories will not be abused for profit. We could share personal information to find love and companionship. We could use Alexas or Cortanas knowing no one was listening in without our consent.

2. Empowering the FTC

We need a privacy regulator capable of investigating and rooting out corporate deception and manipulation that is subtler, yet far more insidious than lying on a privacy policy. An information fiduciary approach would help, but as Daniel Solove and Woodrow Hartzog have shown, the FTC already has a long track record of regulating deceptive practices that erode our privacy.[263] To guard against social robots creating false veneers—so-called "Wizard-of-Oz" setups and misleadingly operated devices, for

259. *Id.* at 2263 (Gorsuch, J., dissenting).

260. Data Care Act, S. 3744, 115th Cong. (2018), https://www.congress.gov/bill/115th-congress/senate-bill/3744 [https://perma.cc/4PWJ-B4RV].

261. *See* WALDMAN, PRIVACY AS TRUST, *supra* note 24, at 86–87; *see also* Balkin, *supra* note 254, at 1222–23; Jack M. Balkin & Jonathan Zittrain, *A Grand Bargain to Make Tech Companies Trustworthy*, ATLANTIC (Oct. 3, 2016), https://www.theatlantic.com/technology/archive/2016/10/information-fiduciary/502346/ [https://perma.cc/QCG 7-7Y4Z].

262. *See* Balkin & Zittrain, *supra* note 261.

263. *See* Solove & Hartzog, *supra* note 195, at 631–45. I am skeptical of the FTC becoming a robust privacy regulator for systemic, structural, and political reasons. That discussion is beyond the scope of this paper, but will be discussed in future research.

example[264]—Hartzog notes that the FTC has already challenged deceptive demonstrations that present products as better than they actually are.[265] And to address the ways in which robots gather information while evading traditional notice modalities, Hartzog notes that the FTC should take Ryan Calo's suggestion to shift to requiring "visceral" notices that leverage "a consumer's very experience of a product or service to warn or inform."[266]

In short, the FTC can pay additional attention to predatory, unfair, or deceptive design that elicits disclosure while eroding trust. Its steps in that direction are promising, but insufficient. For example, in *Sony BMG*, the FTC challenged a design decision that automatically installed digital rights management software on customers' computers without notice and without a reasonable simple way of removing it. [267] In *Frostwire*, the FTC took action when a program automatically designated some files available for public sharing, suggesting an interest in default settings.[268] Although both of these cases involve design, the FTC is still slow to give up its reliance on broken promises litigation. In *Sony BMG*, for instance, the phrase "Respondent has failed to disclose, or has failed to disclose adequately, that" prefaces two substantive allegations against the company; [269] the third simply alleged that deceptive software was "installed on consumers' computers without adequate notification and consent."[270] *Frostwire* more squarely addressed design,[271] but even there the FTC noted that the company failed to disclose what its software would actually do.[272] Emphasizing the role trust plays in user privacy and safety means recognizing that certain design choices can be deceptive even with conspicuous notice. The FTC isn't there yet.

264. *See* Hartzog, *Unfair and Deceptive Robots*, *supra* note 160, at 793–94.

265. *See* Solove & Hartzog, *supra* note 195, at 630–33; *see also* Volvo N.A. Corp., 115 F.T.C. 87 (1992) (finding an advertisement that depicted a monster truck crushing all rival cars except a Volvo deceptive because the Volvo in the ad, unlike Volvo's sold to customers, had been reinforced while the other cars' rooves were weakened).

266. *See* Hartzog, *Unfair and Deceptive Robots*, *supra* note 160, at 818 (quoting Calo, *Against Notice Skepticism*, *supra* note 162, at 1030).

267. Complaint, Sony BMG Music Entm't, F.T.C. File No. 062 3019, No. C-4195 (F.T.C. June 28, 2007), http://www.ftc.gov/sites/default/files/documents/cases/2007/01/070130cmp0623019.pdf [https://perma.cc/Q99B-PP39] [hereinafter Sony BMG Complaint].

268. Complaint for Permanent Injunction and Other Equitable Relief at ¶ 13, FTC v. Frostwire, LLC, No. 11-cv-23643 (S.D. Fla. Oct. 12, 2011), https://www.ftc.gov/sites/default/files/documents/cases/2011/10/111011frostwirecmpt.pdf [https://perma.cc/BYR5-3KKX] [hereinafter Frostwire Complaint].

269. *See* Sony BMG Complaint, *supra* note 267, at ¶¶ 17–18.

270. *Id.* at ¶ 19.

271. *See* Frostwire Complaint, *supra* note 268, at ¶¶ 25, 30, 32, 35–37, 41 (alleging unfair design independent of a failure to notify).

272. *Id.* at ¶ 38.

The FTC also needs the ability to write rules to clarify its authority. The purpose of agency rulemaking is to specify vague statutory requirements, offering clear notice as to what the law requires, an opportunity to participate in public governance, and a comprehensive resolution of questions facing large numbers of persons and businesses.[273] However, the FTC is limited by the "procedurally burdensome" process of Magnuson-Moss rulemaking,[274] which requires the FTC to conduct industry-wide investigations, prepare reports, propose rules, engage in a series of public hearings, and consider other alternatives.[275] The process is so difficult that the FTC has not engaged in this type of rulemaking in thirty-seven years.[276] This lack of rulemaking authority ensures that, without more, privacy regulation from the FTC will remain vague; the only other way to discern what the FTC means by a specific term or phrase is to turn to its previous consent decrees, which is what many practitioners do.[277] But that common law analysis cannot achieve the level of clarity rulemaking can. If applied to Section 5 of the FTC Act, which only prohibits "unfair and deceptive" practices, and any other privacy statute, rulemaking could help clarify what companies have to do to ensure trust-enforcing design.

273. *See* William S. Jordan III, *Ossification Revisited: Does Arbitrary and Capricious Review Significantly Interfere with Agency Ability to Achieve Regulatory Goals Through Information Rulemaking?*, 94 NW. U. L. REV. 393, 394 (2000).

274. Solove & Hartzog, *supra* note 195, at 620. In his comprehensive analysis of the history and development of the FTC, Chris Hoofnagle notes that after several years of rulemaking authority, the Federal Trade Commission Improvement Act of 1980 placed additional procedural hurdles in the FTC's rule-making powers. For example, the Act introduced direct Congressional oversight. And the law explicitly prohibited the FTC from using funds for three years "for the purpose of initiating any new rulemaking proceeding . . . which prohibits or otherwise regulates any commercial advertising." *See* CHRIS J. HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND POLICY 65 (2016) (citing Pub. L. No. 96-252, 94 Stat. 474 (1980)). Even though that ban only explicitly applied for three years, the procedural burdens remain and the Act did much "political and psychological damage to the Agency." *Id.* Notably, the FTC does have general rulemaking authority under the Children's Online Privacy Protection Act and the Gramm-Leach-Bliley Act. *See A Brief Overview of the Federal Trade Commission's Investigative and Law Enforcement Authority*, at app. C, FTC (July 2008), http://www.ftc.gov/about-ftc/what-we-do/enforcement-authority [https://perma.cc/824H-FNT2] ("Special Statutes that mandate or authorize Commission rulemakings either antitrust and/or consumer protection related . . . include the Graham-Leach-Bliley Act . . . [and] COPPA . . . .").

275. FED. TRADE COMM'N, RULEMAKING: OPERATING MANUAL, CHAPTER SEVEN, https://www.ftc.gov/sites/default/files/attachments/ftc-administrative-staff-manuals/ch07rulemaking.pdf [https://perma.cc/PAA2-ZGC P] (describing rulemaking procedures).

276. *See* Solove & Hartzog, *supra* note 195, at 620 n.176.

277. *Id.* at 585 ("Those involved with helping businesses comply with privacy law--from chief privacy officers to inside counsel to outside counsel--parse and analyze the FTC's settlement agreements, reports, and activities as if they were pronouncements by the Chairman of the Federal Reserve.").

3. Privacy by Design

We can also build on the model of the GDPR and codify privacy by design as a legal mandate.[278] Privacy by design is the notion that privacy should be part of the design process for new technologies rather than tacked on at the end as an afterthought. It is meant to proactively protect consumer privacy and reduce the risks consumers face when using data collection tools while making privacy a priority inside technology companies.[279]

But beyond this general understanding, there remains considerable uncertainty as to what privacy by design will mean in practice. The GDPR's formulation is so broad it is almost devoid of meaning, only requiring data collectors take technical and organizational steps to implement the data protection principles listed elsewhere in the GDPR.[280] The academic literature includes no fewer than seven other definitions of privacy by design, none of which gives companies, consumers, and regulators sufficient notice as to the law's practical requirements.[281] Elsewhere, I have argued that privacy by design can transition to privacy's law of design by learning from the law of products liability for design defects. It would then require technology companies to, throughout a product's lifecycle, balance the products' benefits to consumers against their foreseeable privacy risks and only place in commerce those products that achieve reasonably similar consumer benefit with the least privacy risk. This duty would include the responsibility to inform users, throughout the lifecycle of products, of how the products collect and process data and of all foreseeable privacy risks in a manner that adequately and comprehensibly conveys those risks to an ordinary user.[282] By providing a specific legal frame in which designers can innovate new technologies, this formulation of privacy by design would both provide for trust and encourage the organizational changes necessary to make privacy a priority in the ethos and practice of a company.

278. *See* GDPR, *supra* note 220, art. 25, at 48.

279. *See* ANN CAVOUKIAN, PRIVACY BY DESIGN 3 (2009), https://www.ipc.on.ca/wp-content/uploads/Resources/PrivacybyDesignBook.pdf [https://perma.cc/Q4CW-C6YU]; *see also* ANN CAVOUKIAN, PRIVACY BY DESIGN: THE SEVEN FOUNDATIONAL PRINCIPLES (2009), https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf [https://perma.cc/5M4T-FQ58].

280. *See* GDPR, *supra* note 220, art. 25, at 48 ("the controller shall . . .. . . implement appropriate technical and organisational measures . . . which are designed to implement data-protection principles . . . in an effective manner").

281. *See* Waldman, *Privacy's Law of Design*, *supra* note 26, at 1253–56.

282. *Id.* at 1285.

4. Reform to Section 230

More immediate relief can come from modest reform to the broad immunity of Section 230.[283] Danielle Citron and Benjamin Wittes have proposed that Section 230 immunity be extended only to "good Samaritans," or "providers or users engaged in good faith efforts to restrict illegal activity."[284] This is a good idea, and one that would incent platforms to protect their users and foster the kind of trust necessary for social interaction while preserving robust speech online and protecting companies from being buried under lawsuits for honest mistakes.

Protecting only those platforms that act in good faith is, in fact, what Congress intended from the beginning.[285] Section 230(c)(2),[286] which immunizes platforms from suits related to any "action voluntarily taken in good faith to restrict access to . . . material that . . . [is] harassing or otherwise objectionable," is actually called the "Good Samaritan" provision.[287] It was meant to incent trust-enhancing design, not create social spaces that look like the Wild West. That the federal courts interpreting Section 230 turned away from Congressional intent and the plain language is an accident of history: a bug, rather than a feature, of the law.

This kind of limitation on platform immunity would protect platforms that earnestly try to filter out harassing content but make honest mistakes. It would, however, constrain those platforms that ignore their users' consistent complaints about racism, harassment, revenge pornography, and other behaviors that violate terms of service and make technosocial spaces unsafe. It would, in short, buttress content moderation designed for user safety and trust.

CONCLUSION

This Article began with a problem: technosocial spaces can be unsafe and privacy invasive. But it also saw an opportunity: to learn by analogy from offline social spaces—Alcoholics Anonymous, teams of coworkers, and attorney-client relationships—that facilitate socially beneficial disclosures by counterbalancing norms of disclosure with equally powerful norms of trust that are both endogenously designed and exogenously

283. *See supra* notes 210–219 and accompanying text.
284. Citron & Wittes, *supra* note 213, at 416.
285. For a more comprehensive review of the "origin story," please see *id.* at 404–411.
286. 47 U.S.C. § 230(c)(2).
287. Batzel v. Smith, 333 F.3d 1018, 1027 (9th Cir. 2003) (citing 47 U.S.C. § 230 (b)) (Section 230 was meant "(1) to promote the continued development of the Internet and other interactive computer services and other interactive media; (2) to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by Federal or State regulation").

supported by law. With those analogies in mind, the Article proposed several specific ways trust can be designed in to technosocial spaces using code and four legal levers—fiduciary law, consumer protection, products liability, and intermediary liability—that could be leveraged to support trust and safety.

This article is also situated in a growing scholarly literature on trust and the law. Although its focus was on how trust and law can affect technologically mediated social life, trust also plays important roles in our relationships to government, law enforcement, health care companies, landlords, employers, and more. This research agenda is, therefore, ripe for growth.

In the end, no social spaces, online or offline, can always be safe. Life involves risk, and so do disclosures, social networking, online dating, and the conveniences of modern life. But privacy and safety remain relevant. Privacy, expectations of confidentiality and discretion, and relief from hate and harassment are all necessary for identity formation, intellectual freedom, and equality. As such, they are essential to a well-functioning democracy, increasingly under threat today. Design and law can play guiding and expressive roles in support of these goals.