ONLINE HARASSMENT AND INTERMEDIARY IMMUNITY

WILLIAM MCGEVERAN*

Throughout much of *Hate Crimes in Cyberspace*,[1] Danielle Citron sounds a clarion call for reform to prevent abusive behavior online, especially toward women. As she demonstrates vividly with her case studies, this is a civil rights crisis. And she has lots of ideas to combat it. When the discussion turns to liability for internet intermediaries, however, Citron's tone becomes more cautious and subdued. This shows good judgment: there isn't much reason to hold back where the only opposing forces are apathy and misogyny, but this particular issue involves meaningful interests in free expression. That said, there may now be scope to do more than Citron proposes.

We can learn about our options from two previous situations where the law has held online intermediaries responsible for content: the notice-and-takedown regime of the Digital Millennium Copyright Act (DMCA) here in the United States, and the emerging "right to be forgotten" in the European Union. Neither of these laws has "broken the internet," to use Citron's phrase, although I shall suggest that each has shortcomings to be avoided in designing any remedy for online harassment.

As Citron explains, the federal law usually known as "§ 230" immunizes an "interactive computer service" from liability for most claims related to user-generated content, including those typically raised by harassment or nonconsensual pornography.[2] The only exemptions to this safe harbor cover federal criminal and eavesdropping law and intellectual property infringement. As a result, many victims of online attacks have no realistic recourse. Citron bemoans the problem, but is extremely careful not to embrace an overly broad solution. She criticizes a proposal supported by most state attorneys general to include all state criminal law among the exemptions to § 230 immunity, because that would "require online providers to shoulder burdensome legal compliance with countless state criminal laws that have nothing to do with the most troubling uses of online platforms . . . ."[3] Instead, Citron suggests a much more tailored amendment that would apply only to the sites she calls "cyber cesspools"—those that principally host harassment or nonconsensual pornography.

* Associate Professor and Solly Robins Distinguished Research Fellow, University of Minnesota Law School.

[1] DANIELLE CITRON, HATE CRIMES IN CYBERSPACE (2014).

[2] 47 U.S.C. § 230(c). Because the provision broadly prohibits treating an intermediary as a publisher, platforms that host user-generated content avoid liability for privacy torts, defamation, and many other claims.

[3] CITRON, *supra* note 1, at 177.

Citron acknowledges that even this modest change might be too much for some technologists and civil libertarians whose fervent devotion to § 230 verges on the worshipful. They don't much care for the intellectual property exemption, and they certainly oppose any further diminution of the safe harbor. At the same time, other activists will be frustrated that Citron offers a limited response to the problems she documents. They may think her reform should reach sites like AutoAdmit or 4chan that host plentiful harassing content, not to mention malicious posts made through mainstream social media platforms.

I admire Citron for listening to all sides and adopting a nuanced position that recognizes the thin line separating a cyber cesspool from a public-spirited open forum like Reddit or Wikipedia. But I am also pleased that in some ways her defensiveness is already a tad dated. Attitudes are changing—partly because of scholarship by Citron and others,[4] partly because of the activism of organizations like the Cyber Civil Rights Initiative and Without My Consent, and partly because of outrage at incidents like the theft of private nude photos of Jennifer Lawrence and other celebrities. So, without denying the real and important speech interests sometimes implicated, maybe it is possible to think slightly more ambitiously. The DMCA and right to be forgotten offer guidance.

Under the DMCA, an intermediary must respond to proper notice from a copyright holder identifying alleged infringement by a user. The ordinary response is to remove the content. The person who originally posted it may respond with a counter-notice, essentially providing the other side of the story. In that case, if the copyright holder does not sue over the matter, the intermediary is free to restore the content.[5]

This arrangement has shortcomings. Allowing one objection to block access to speech may enable a heckler's veto, as Citron notes. Recent headlines lend credence to this concern. The NFL, relying on an extremely broad understanding of its copyright, just forced Twitter to disable accounts that post brief GIFs of highlights taken from game footage, many of which surely qualify as fair use.[6] That said, the overwhelming majority of takedown requests involve blatant unlawful copying. Automated systems identify many of them. And some protections exist against the heckler's veto, including the statutory provision for counter-notice by the person whose speech may be curtailed. In addition, the Ninth Circuit just ruled that a copyright holder is required to consider possible fair use arguments before sending a takedown request.[7] The problems come

[4] *See, e.g.*, Ann Bartow, *Internet Defamation as Profit Center: The Monetization of Online Harassment*, 32 HARV. J.L. & GENDER 383 (2009); Danielle Keats Citron & Mary Anne Franks, *Criminalizing Revenge Porn*, 49 WAKE FOREST L. REV. 345 (2014); Mary Anne Franks, *Sexual Harassment 2.0*, 71 MD. L. REV. 655 (2012).

[5] *See* 17 U.S.C. § 512.

[6] *See* Mathew Ingram, *Here's Why Deadspin Is Right, and the NFL and Twitter Are Wrong*, FORTUNE.COM (Oct. 13, 2015), http://fortune.com/2015/10/13/deadspin-nfl-twitter/.

[7] *See* Lenz v. Universal Music Corp., No. 13-16106 (9th Cir. Sept. 14, 2015), *available at* http://cdn.ca9.uscourts.gov/datastore/opinions/2015/09/14/13-16106.pdf.

about because the protections are not always effective in practice, and because stifling even a small number of noninfringing uses interferes with robust public discourse and individual interests in free expression. Furthermore, fair use is a complex and fact-sensitive doctrine, making it more difficult to distinguish the meritorious takedowns from the overly broad claims and the meddling hecklers.

The "right to be forgotten" under European data protection law provides another instructive example. Forms of this right have existed in European law for decades.[8] For that matter, American law allows individuals to demand deletion of personal data under certain circumstances.[9] The recent European court ruling in Google Spain SL v. AEPD went much further, however, imposing a duty on a search engine to remove links leading to personal data that is "inadequate, irrelevant or no longer relevant, or excessive in relation to the purposes" of data processing.[10] The ruling tries to give European citizens some measure of control over old and potentially invasive information, much as Jonathan Zittrain envisioned when he contemplated a form of "reputational bankruptcy."[11]

The Google Spain rule imposes costly obligations on Google—and all search engines, and presumably many other types of intermediaries. Google and others responded by fielding requests to remove links and reaching individualized decisions about each one. Furthermore, determinations under the right to be forgotten are highly complex. The Google Spainruling provides hardly any direction for reaching them, leaving vast discretion to intermediaries and further increasing compliance costs.[12] Yet the number of these requests does not approach the number of copyright takedown demands: while Google has now considered over 1.1 million URLs for deletion under the European right to be

[8] *See, e.g.*, Directive 95/46/EC (1995), Art. 12(3) (providing individuals the right to "the rectification, erasure or blocking of data the processing of which does not comply with the provisions of this Directive, in particular because of the incomplete or inaccurate nature of the data").

[9] *See, e.g.*, Fair Credit Reporting Act, 15 U.S.C. § 1681i(a)(5)(A) (requiring the deletion of certain information in credit reports if it is "found to be inaccurate or incomplete or cannot be verified"); Children's Online Privacy Protection Rule, 16 C.F.R. 312.6(a)(2) (allowing parents to demand the deletion of data collected online about their children under the age of 13).

[10] ECLI:EU:C:2014:317 (CJEU).

[11] *See* JONATHAN L. ZITTRAIN, THE FUTURE OF THE INTERNET—AND HOW TO STOP IT 228-29 (2008); *see generally* VIKTOR MAYER-SCHÖNBERGER, DELETE: THE VIRTUE OF FORGETTING IN THE DIGITAL AGE (2009).

[12] *See* Meg Leta Jones, *Forgetting Made (Too) Easy*, 58 COMM. OF THE ACM 34, 35 (2015) ("The tragic news is that Google (and other data controllers) has been handed a gavel and ordered to decide what the right to be forgotten is without any guidance as to which interests should trump others, when, and why.").

forgotten since the Google Spain decision, it received 50 million takedowns from copyright holders last month.[13]

Neither the DMCA nor the right to be forgotten has proven disastrous, but both systems have flaws. From them, we learn that clear standards, limited need for intermediaries to make tough judgment calls, and a right of reply are all important. Applying their lessons realistically and keeping the actual threat to speech interests in view, it is possible to go further than Citron does in her book.

For example, Citron proposes criminalizing revenge porn (appropriately defined to avoid clashing with the public interest).[14] If this became a federal crime, of course, it would fall outside of the current form of § 230—a neat solution. But even if not, revenge porn as she defines it could anchor a reasonable and balanced amendment to § 230. What if intermediaries had to accept notice of revenge porn allegations just as they now do for copyright, and act on them once received? Surely the number of complaints would be only a subset of those now covered by the right to be forgotten. Moreover, nude or sexual images as a category are much easier to identify and isolate than either copyright infringement or the wide range of materials covered under Google Spain; automated software does a pretty good job of spotting it already. And intermediaries need not be required to do anything until and unless they receive a notice. Certainly, frivolous complaints might be lodged, but counter-notice could be made available. And finally: what are the likely situations (not the hypothetical ones) where a temporary takedown of a sexual image throttles speech? How does that remote potential compare to the demonstrable harm Citron documents from such harassment? We already regulate pornography in numerous ways, from zoning law to age restrictions. This would only be a minor addition to those, and more clearly justified than some of them.

There is no reason § 230 should offer vastly more recourse to Walt Disney and Marvel Comics than to the women described in the first portion of *Hate Crimes in Cyberspace*. Thanks to the work of Citron and many others, there is reason to think that in the near future, it may not.

[13] *See* GOOGLE, *Transparency Report*, https://www.google.com/transparencyreport/ (last visited Oct. 15, 2015) (copyright numbers cover the period from Sept.15, 2015 to Oct. 15, 2015); *see also* MICROSOFT, *Content Removal Requests Report*, http://www.microsoft.com/about/corporatecitizenship/en-us/transparencyhub/crrr/ (last visited Oct. 15, 2015) (showing requests to remove over 10,000 URLs on basis of right to be forgotten between January and June 2015, and over 24.5 million URLs on basis of copyright).

[14] CITRON, SUPRA note 1, at 145-53.