

2018

Justice Beyond Dispute

Mary Anne Franks

University of Miami School of Law, mafranks@law.miami.edu

Follow this and additional works at: https://repository.law.miami.edu/fac_articles

 Part of the [Dispute Resolution and Arbitration Commons](#), [Internet Law Commons](#), and the [Law and Society Commons](#)

Recommended Citation

Mary Anne Franks, *Justice Beyond Dispute*, 131 *Harv. L. Rev.* 1374 (2018).

This Book Review is brought to you for free and open access by the Faculty and Deans at University of Miami School of Law Institutional Repository. It has been accepted for inclusion in Articles by an authorized administrator of University of Miami School of Law Institutional Repository. For more information, please contact library@law.miami.edu.

BOOK REVIEW

JUSTICE BEYOND DISPUTE

DIGITAL JUSTICE: TECHNOLOGY AND THE INTERNET OF DISPUTES.
By Ethan Katsh and Orna Rabinovich-Einy. New York, N.Y.: Oxford
University Press. 2017. Pp. xi, 910. \$24.95.

*Reviewed by Mary Anne Franks**

I. INTRODUCTION

Jiranuch Triratana was watching her brother scroll through his Facebook feed on April 24, 2017, when they came upon a startling live-stream broadcast. It was from Jiranuch's boyfriend, Wuttisan Wongtalay.¹ He was filming himself and the couple's eleven-month-old daughter, Natalie, from the roof of a building.² There was a rope tied around Natalie's neck.³ As Jiranuch and her brother watched in horror, Wongtalay dropped the infant off the side of the building.⁴ Jiranuch alerted the police, who found the lifeless bodies of Natalie and her father hanging from ropes off the side of an abandoned hotel in Phuket a few hours later.⁵ Wongtalay had killed his daughter and then himself.⁶ Wongtalay's smartphone was propped up against a nearby wall.⁷

The two Facebook Live videos showing Wongtalay murdering his daughter were available on the platform for roughly twenty-four hours.⁸ Before Facebook removed them, one video had been viewed 112,000 times and the other 258,000 times.⁹ Both had been uploaded to YouTube.¹⁰ Facebook responded with a statement that called the killing

* Professor of Law, University of Miami School of Law.

¹ Agence France-Presse, *Thai Mother Saw Daughter Being Killed on Facebook Live*, THE GUARDIAN (Apr. 27, 2017, 2:49 AM), <https://www.theguardian.com/world/2017/apr/27/thai-mother-watched-daughter-being-killed-on-facebook-live> [<https://perma.cc/KFK6-5UXK>].

² Patpicha Tanakasempipat & Panarat Thepgumpanat, *Thai Man Broadcasts Baby Daughter's Murder Live on Facebook*, REUTERS (Apr. 25, 2017, 7:47 AM), <https://www.reuters.com/article/us-thailand-facebook-murder-idUSKBN17R1DG> [<https://perma.cc/G54T-G2S7>].

³ Agence France-Presse, *supra* note 1.

⁴ Tanakasempipat & Thepgumpanat, *supra* note 2.

⁵ Agence France-Presse, *supra* note 1.

⁶ *Id.*

⁷ Shira Rubin, *Thai Man Kills Baby Daughter, Then Himself, Live on Facebook*, VOCATIV (Apr. 25, 2017, 11:05 AM), <http://www.vocativ.com/424039/thai-man-kills-daughter-facebook-live/> [<https://perma.cc/M9UP-GBZ2>].

⁸ Tanakasempipat & Thepgumpanat, *supra* note 2.

⁹ *Id.*

¹⁰ Jessica Guynn, *Father Livestreams Killing of Infant Daughter on Facebook Live*, USA TODAY (Apr. 25, 2017, 12:59 PM), <https://www.usatoday.com/story/tech/news/2017/04/25/father-livestreams-killing-infant-daughter-facebook-live/100884906/> [<https://perma.cc/MLX8-KNSF>].

“an appalling incident” and asserted that “[t]here is absolutely no place for acts of this kind on Facebook.”¹¹

But acts of this kind have found a place on Facebook with increasing frequency. At least sixty violent incidents have been broadcast on Facebook Live, the company’s live-streaming service, between its launch in December 2015 and April 2017.¹² The incidents include “shootings, rapes, murders, child abuse, torture, suicides, and attempted suicides.”¹³ In January 2017, a Facebook Live video of four people in Chicago beating a bound and gagged mentally disabled teenager, at one point cutting into his scalp with a knife,¹⁴ was left up for “at least 23 hours and was viewed more than 16,000 times before Facebook’s reviewers intervened.”¹⁵ In March 2017, at least forty people watched the gang rape of a fifteen-year-old girl on Facebook Live, none of whom called the police.¹⁶ In May 2017, a thirty-three-year-old man who had recently been arrested for trying to kill his ex-girlfriend turned on Facebook Live just before he doused himself in kerosene, ran into the bar where she worked, and set himself on fire in front of her.¹⁷

One of the highest-profile cases of violence involving Facebook Live unfolded in Cleveland on Easter Sunday in 2017. That day, Steve Stephens posted a video to Facebook declaring his intention to commit murder.¹⁸ A few minutes later, he posted a video of himself approaching

¹¹ *Id.*

¹² Edgar Alvarez, *Murders, Suicides and Rapes: Facebook’s Major Video Problem*, ENGADGET (Apr. 18, 2017), <https://www.engadget.com/2017/04/18/facebook-video-steve-stephens/> [<https://perma.cc/JM78-YVGQ>]; see also Alex Kantrowitz, *Violence on Facebook Live Is Worse than You Thought*, BUZZFEED (June 16, 2017, 10:17 AM), <https://www.buzzfeed.com/alexkantrowitz/heres-how-bad-facebook-lives-violence-problem-is> [<https://perma.cc/5A3Y-GJBG>].

¹³ Kantrowitz, *supra* note 12.

¹⁴ Steve Schmadeke et al., *Four Ordered Held Without Bail in Alleged Hate-Crime Attack Streamed Live on Facebook*, CHI. TRIB. (Jan. 6, 2017, 8:50 PM), <http://www.chicagotribune.com/news/local/breaking/ct-facebook-live-attack-charges-met-20170106-story.html> [<https://perma.cc/BGM7-ELSC>].

¹⁵ Deepa Seetharaman, *Facebook, Rushing into Live Video, Wasn’t Ready for Its Dark Side*, WALL ST. J. (Mar. 6, 2017, 2:05 PM), <https://www.wsj.com/articles/in-rush-to-live-video-facebook-moved-fast-and-broke-things-1488821247> [<https://perma.cc/S7YA-D6LU>].

¹⁶ Phil McCausland & Associated Press, *Teen Who Was Gang Raped on Facebook Live Is Receiving Threats, Mom Says*, NBC NEWS (Mar. 22, 2017, 6:19 PM), <http://www.nbcnews.com/news/us-news/teen-who-was-gang-raped-facebook-live-receiving-threats-mom-n737356> [<https://perma.cc/8BHM-AZGF>].

¹⁷ Shehab Khan, *Man Dies After Setting Himself on Fire During Facebook Live Stream*, THE INDEPENDENT (May 15, 2017, 9:08 PM), <http://www.independent.co.uk/news/world/americas/facebook-live-man-dies-sets-self-on-fire-suicide-stream-a7737621.html> [<https://perma.cc/L6R2-HK9A>]; see also Kristine Phillips & Peter Holley, *He Doused Himself with Kerosene on Facebook Live — Then Ran into a Bar in Flames*, WASH. POST (May 15, 2017), <https://www.washingtonpost.com/news/true-crime/wp/2017/05/15/he-doused-himself-with-kerosene-on-facebook-live-then-ran-into-a-bar-in-flames/> [<https://perma.cc/HJ7H-LANE>].

¹⁸ Elizabeth Dwoskin & Craig Timberg, *Facebook Wanted “Visceral” Live Video. It’s Getting Live-Streaming Killers and Suicides*, WASH. POST (Apr. 17, 2017), <https://www.washingtonpost.com/business/technology/facebook-wanted-visceral-live-video-its-getting-suicides->

an elderly man, later identified as Robert Godwin, Sr., and asking him to say a woman's name.¹⁹ After Godwin, who appeared confused by the request, said the name,²⁰ Stephens told him, "She's the reason that this is about to happen to you."²¹ Stephens then fatally shot Godwin in the head.²² A few minutes later, Stephens used Facebook Live to broadcast himself confessing to murder.²³ According to Facebook, Stephens's Facebook account was disabled a few hours after he posted the first video. By then, the video of the murder had been posted across the platform and on other social media sites.²⁴ Within hours of the killing, one post of the video had been viewed 1.6 million times.²⁵

The outcry over Facebook Live videos of graphic murders, rapes, and beatings has been passionate and sustained. Godwin's grandson begged people to "show some respect" and stop sharing the video of his grandfather's murder,²⁶ highlighting the agony inflicted on family members when such tragedies go viral. Psychologists have observed that being repeatedly exposed to acts of violence can lead to desensitization or to "secondary trauma," which suggests that watching real-life videos of murder, rape, and other violence can cause lasting psychological damage.²⁷ This danger is particularly acute for social media content moderators who are tasked with viewing violent, disturbing videos for hours on end.²⁸ Experts warn that live-streaming technology like Facebook Live "serves in some cases as an impetus to some people" to commit acts of violence.²⁹ Criminals obtaining notoriety through live streaming, criminologists warn, can attract attention-seekers and inspire copycat

and-live-streaming-killers/2017/04/17/a6705662-239c-11e7-afb3-faff034e2de_story.html [https://perma.cc/64CD-7PKM].

¹⁹ Jane Morice, *Facebook Killer Chooses Victim at Random, Laughs About Killing in Videos*, CLEVELAND.COM (Apr. 16, 2017), http://www.cleveland.com/metro/index.ssf/2017/04/accused_facebook_live_killer_c.html [https://perma.cc/R7LL-FMNT].

²⁰ *Id.*

²¹ Melissa Chan, *What to Know About Cleveland Facebook Murder Suspect Steve Stephens*, TIME (Apr. 17, 2017, 3:27 PM), <http://time.com/4742204/steve-stephens-cleveland-shooting-facebook/> [https://perma.cc/H9EQ-72GM].

²² *Id.*

²³ Dwoskin & Timberg, *supra* note 18.

²⁴ *Id.*

²⁵ Courtney Astolfi, *Social Media Users Urge Against Sharing Facebook Video of Cleveland Killing*, CLEVELAND.COM (Apr. 16, 2017), http://www.cleveland.com/metro/index.ssf/2017/04/social_media_users_urge_agains.html [https://perma.cc/HYU9-YVRX].

²⁶ *Id.*

²⁷ Olivia Solon, *Facebook Is Hiring Moderators. But Is the Job Too Gruesome to Handle?*, THE GUARDIAN (May 4, 2017, 5:00 AM), <https://www.theguardian.com/technology/2017/may/04/facebook-content-moderators-ptsd-psychological-dangers> [https://perma.cc/B8B5-7MYD].

²⁸ *Id.*

²⁹ See Rod Chester, *Facebook to Hire 3000 People to Stop Violence and Suicide in Streaming Live Video*, NEWS.COM.AU (May 4, 2017, 7:32 AM), <http://www.news.com.au/technology/online/social/facebook-to-hire-3000-people-to-spot-and-stop-violence-and-suicide-in-streaming-live-video/news-story/8202db56d8d9764d4ee85c179ef72e15> [https://perma.cc/M7UH-M3B8].

violence.³⁰ The longer a video remains on a platform, the larger the audience it will reach and the more likely it is that it will migrate to other sites. For all of these reasons, critics have suggested that Facebook should act faster to take down violent live streams,³¹ impose a time delay for broadcasts similar to the seven-second delay television networks use for broadcasting live events,³² or get rid of its live-streaming service altogether.³³

Facebook founder Mark Zuckerberg, speaking at the company's annual developer conference two days after the video of Godwin's murder went viral, made only a brief reference to the incident, saying "we will keep doing all we can to prevent tragedies like this from happening" before moving on to discuss the company's plans for augmented-reality technology.³⁴ The criticism over Facebook Live did not abate, intensifying after the video of Wongtalay's murder of his daughter went viral a little more than a week later.³⁵ On May 3, 2017, Zuckerberg offered a more substantive response to the outcry, stating on the Facebook website that the company would add 3000 moderators "to review the millions of reports we get every week, and improve the process for doing it quickly."³⁶

Given that live-stream content must still first be flagged by a user before a moderator can review it, the only way that removals can be done more quickly and effectively is for more users to watch and flag more traumatizing content. Adding more moderators may speed up the process at the secondary stage, but ultimately the burden is — that is, remains — on users. Such a response does nothing to address the fundamental problem Facebook created with its live-stream service. "Pure" live streaming — zero-delay broadcasting — is by design impossible to moderate in any meaningful way. A system that allows content to be broadcasted unless and until flagged by a user and reviewed by a moderator is a system that will inevitably put rapes, suicides, and murders into public view. Videos of such gruesome acts only need to be available for a few seconds to go viral, at which point removal by the original

³⁰ Gillian Mohney, *Murder on Facebook Spotlights Rise of "Performance Crime" Phenomenon on Social Media*, ABC NEWS (Apr. 18, 2017, 4:47 PM), <http://abcnews.go.com/US/murder-facebook-spotlights-rise-performance-crime-phenomenon-social/story?id=46862306> [<https://perma.cc/K2MN-L3R6>].

³¹ See, e.g., Kantrowitz, *supra* note 12.

³² See, e.g., Alvarez, *supra* note 12.

³³ See, e.g., Seetharaman, *supra* note 15.

³⁴ Nancy Dillon, *Mark Zuckerberg Glosses over Steve Stephens' Viral Facebook Murder in Conference Talk*, N.Y. DAILY NEWS (Apr. 18, 2017, 3:33 PM), <http://www.nydailynews.com/news/national/facebook-lot-work-live-murder-video-ceo-article-1.3069887> [<https://perma.cc/RU23-TA3M>].

³⁵ Samuel Gibbs, *Facebook Under Pressure After Man Livestreams Killing of His Daughter*, THE GUARDIAN (Apr. 25, 2017, 11:38 AM), <https://www.theguardian.com/technology/2017/apr/25/facebook-thailand-man-livestreams-killing-daughter> [<https://perma.cc/P4U5-MZXH>].

³⁶ Mark Zuckerberg, FACEBOOK (May 3, 2017), <https://www.facebook.com/zuck/posts/10103695315624661> [<https://perma.cc/C28V-3DYU>].

platform will have limited impact. Increasing the number of people assigned to a futile task does not make the task any less futile; it merely increases the number of people who will be subjected to traumatizing content.

Professors Ethan Katsh and Orna Rabinovich-Einy's new book, *Digital Justice: Technology and the Internet of Disputes*, does not discuss Facebook Live, but it does praise what the authors call Facebook's efforts to meet the problem of "anti-social media" (p. 109). Facebook's implementation of a dispute resolution system that allows users to work out conflicts among themselves demonstrates, according to the authors, an "advanced understanding of the role of dispute systems design" (p. 121). This dispute system allows users to communicate directly with each other about behavior they find offensive, "explaining how it makes them feel and what actions they would like to be taken by their counterpart" (p. 120). While the authors acknowledge that the process can give rise to other disputes and has the potential to be abused, it nonetheless "reflects a recognition of the significance of dispute resolution processes that create a space for users to discuss problems, feelings, and desired outcomes" (p. 121). In doing so, Facebook eschews a decisionmaking role and "allows instead for direct user-to-user negotiation" (p. 121).

Digital Justice is a book about "online dispute resolution," also known as ODR, and is primarily focused on disputes among parties who have voluntarily entered into relationships that were intended to serve common interests but have broken down at some point (p. 113). Many transactional and commercial disputes can be characterized in this way, as can relational disputes involving parties of roughly equal standing: e-commerce transactions between buyers and sellers, medical record corrections between patients and hospitals, child custody proceedings between the parents, and workplace disagreements. In pro-social disputes, the interests of corporate or institutional entities and individual parties usually align in some way: online commercial enterprises want customers to continue using their service, hospitals are invested in the accuracy of patient files, the parents of a child want to agree on custody arrangements, employers prefer peaceful workplaces. In these kinds of pro-social interactions, the parties are united by a shared appetite for dispute resolution (p. 113). When disputes arise in these interactions, values such as convenience (p. 37), speed (pp. 74-75), user control (pp. 120-21), and a preference for systematic solutions (pp. 34-36) are particularly valuable in achieving resolution. Most importantly, online dispute resolution emphasizes technology's role as a "Fourth Party" that has the capacity to enhance all of these virtues (pp. 37-38).

The most intense and damaging social media conflicts, however, can be described as aggressively anti-social. As Katsh and Rabinovich-Einy observe in Chapter 5, "The Challenge of Social and Anti-Social Media,"

social media conflicts can be more challenging than other kinds of conflicts because they “have almost none of the qualities or systems” that are useful in resolving commercial disputes (p. 114). Transactional disputes often arise out of misunderstandings and accidents that can be corrected relatively easily, whereas social media conflicts often involve intentionally malicious acts (p. 114). What is more, the kind of automated infrastructure often available in transactional disputes, for example, credit card “chargeback” systems, is not available for social media disputes (p. 114).

While noting that significant differences in the nature of transactional disputes and social media disputes make the latter more resistant in some ways to dispute resolution, Katsh and Rabinovich-Einy nonetheless endorse ODR as an effective approach to anti-social conflicts as well as pro-social conflicts. Indeed, as the title of their book indicates, they believe that ODR is not merely a useful tool for resolving transactional disputes, but a broad practice that can help produce “digital justice.” But the authors’ discussion of how ODR can and should be applied to anti-social disputes exposes the severe limitations of this vision of justice.

The authors distinguish between commercial or transactional disputes and what they call “relationship disputes.” The primary difference, in their view, is that the latter is more emotional than the former. Katsh and Rabinovich-Einy describe relationship disputes as often involving “actual friends” or at least “people who know each other,” who have “done something online to anger or embarrass the other party” (p. 115). In commercial disputes, which often involve monetary loss, “restitution is often sufficient to achieve a resolution” (p. 115). By contrast, “relationship disputes often require attention to emotions,” especially anger (p. 115).

This characterization, which conjures up short-lived tiffs between impulsive teenagers, is a caricature of online conflict. The examples of anti-social conflicts that the authors themselves provide make the inappropriateness of this characterization clear. Such a description does not fit death threats, rape threats, defamation, revenge porn (the unauthorized disclosure of sexually explicit imagery), and doxxing (the publication of private information online with the purpose of harassing the target) (p. 118),³⁷ nor does it fit live-streamed suicides, terrorist propaganda, conspiracy theories, and misinformation campaigns. Victims of online abuse withdraw from civic life, develop PTSD and

³⁷ Online abuse is disproportionately aimed at women, racial and religious minorities, and lesbian, gay, bisexual, and transgender individuals. See Danny O’Brien & Dia Kayyali, *Facing the Challenge of Online Harassment*, ELECTRONIC FRONTIER FOUND. (Jan. 8, 2015), https://www.eff.org/deeplinks/2015/01/facing-challenge-online-harassment#footnote1_81jmuwe [https://perma.cc/8D79-G6J6].

depression, lose jobs, flee their homes, change their names, and commit suicide.³⁸ To characterize social media conflicts as emotional “relationship disputes” stemming from “embarrassing” behavior (p. 115) is neither accurate nor respectful to the victims of such abuse.

Social media disputes present challenges for dispute resolution not because they involve emotion, but because they involve exploitation. Though Katsh and Rabinovich-Einy claim to be concerned with justice, they seem curiously unconcerned by the fact that anti-social online behavior generates benefits and costs that are unequally distributed across society. But the ability of some groups to profit — financially or otherwise — from the misery of other groups should be incompatible with any intelligible concept of justice.

Katsh and Rabinovich-Einy do specify that they are concerned with procedural, as opposed to substantive, justice.³⁹ While procedural justice is a vitally important value, it is also an inherently limited one that presumes the legitimacy of a given legal or social order. If an underlying legal or social order is itself unjust, focusing on fairness within that order is at best nonresponsive and at worst complicit in injustice. A procedural justice focus within a system based on exploitation risks naturalizing and depoliticizing that exploitation.

The goal of *Digital Justice*, according to the authors, is to clarify not only “how technology generates disputes of all types” but also “how technology can be employed to resolve and prevent them” (p. 3). In other words, the authors believe that technology can solve the problems created by technology. While this is a pragmatic and useful approach in many contexts, it is also, in essence, a form of technological determinism. Technological determinists assume that society should conform itself to the dictates of technology rather than the other way around.⁴⁰ Katsh and Rabinovich-Einy’s technological determinism is evident in their embrace of the ODR metaphor of technology as a “Fourth Party” aiding conflict resolution (p. 37). This position fails to adequately recognize the degree to which technology itself is an antagonistic party in many online disputes, and how the powerful entities that currently exert near-monopoly control over this technology are fundamentally compromised with regard to conflict resolution.

Viewing technology primarily as a solution rather than a problem erases the political and cultural values embedded in technological practices. Technological determinists do not question whether any particular technology should have come into existence or should stay in existence; technological advancement is taken as an inevitability to which humans can adapt in more or less efficient ways. Katsh and Rabinovich-Einy’s

³⁸ See DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE 5–12 (2014).

³⁹ The authors state early on: “[W]e use the term ‘justice’ primarily in a procedural sense, much in the same way it has been used by the ‘access to justice’ literature” (p. 3).

⁴⁰ LELIA GREEN, THE INTERNET 8 (2010).

technological determinism leads them to pay inadequate attention to three key asymmetries at play in the most intensely anti-social online interactions: asymmetries of consent, power, and labor. These three asymmetries are not only exploited by malicious users, but also by online “intermediaries” — corporations such as Facebook, Google, and Twitter — to maximize their profits.

Consent: Unlike commercial disputes, online conflicts are often completely one-sided. Stalkers, harassers, Gamergate trolls, and purveyors of revenge porn are not engaged in mutual activities. Rather, they force unwilling targets into destructive interactions. Similarly, people who live stream murders, disseminate terrorist propaganda videos, or spread conspiracy theories impose harmful content on unsuspecting users of social media and internet services.

Power: The perpetrators of online abuse wield more power than do their victims. While perpetrators often remain anonymous, their targets are denied that privilege. Accordingly, perpetrators are shielded from the consequences of their actions while victims are forced to contend with the effects of being exposed.⁴¹ Perpetration often emboldens abusers and attracts supportive fellow travelers, while victimization silences and isolates victims.⁴²

Labor: Online abuse forces victims to act. Filling out complaint forms, filing police reports, running constant internet searches on their names, sending takedown notices — all of these steps take time, energy, and often financial resources. Online abuse effectively compels victims to provide free labor to try to protect themselves. An online abuser can inflict massive damage with a click of a button, whereas his victim may have to commit to hours of distressing, repetitive, and often ultimately fruitless work to mitigate the harm.

Unlike the commercial or institutional actors in pro-social disputes, social media platforms in anti-social disputes often have no incentive to resolve or prevent the conflicts at issue. They may in fact have incentives to ignore or even to aggravate them. This is due in large part to the business model of many social media companies. They do not make money by selling products; they make money by selling ads. Increased engagement with their platforms, whether for pro-social or anti-social purposes, translates into increased profits: “[A]busive posts still bring in considerable ad revenue and the more content that is posted, good or bad, the more ad money goes into their coffers.”⁴³ This can create in-

⁴¹ See, e.g., CITRON, *supra* note 38, at 5–10.

⁴² See *id.* at 5.

⁴³ Kalev Leetaru, *Do Social Media Platforms Really Care About Online Abuse?*, FORBES (Jan. 12, 2017, 11:46 AM), <https://www.forbes.com/sites/kalevleetaru/2017/01/12/do-social-media-platforms-really-care-about-online-abuse/#58ea898345f1> [https://perma.cc/K3DK-BR4R].

centives for platforms to be indifferent to or even encouraging of inequalities of power among users. For some of these platforms, online abuse may be, as the saying goes, “not a bug but a feature.”⁴⁴

This is another way of saying that online intermediaries like Facebook, Google, and Twitter themselves aggravate and exploit asymmetries of consent, power, and labor. Many technology platforms have become so ubiquitous and powerful as to be virtually inescapable. A person who has not chosen to use Google’s search engine is nonetheless subject to Google’s indexing practices. A person who has not joined Facebook or Twitter can nonetheless be targeted by users on those sites. Even when users have voluntarily engaged with a platform, the platform’s power greatly outstrips the individual or even collective power of users. Byzantine terms of service and invisible data collection practices ensure that the platform always has more power over users than users have over the platform. The most successful online services, moreover, are primarily products of users’ free labor. Google does not create the web that it indexes; Facebook and Twitter do not write the posts that they promote. When disputes and controversies arise, the burden is almost always on users to perform the work to address them, whether that is flagging live-streamed murders or reporting revenge porn.

Online abuse is rarely extricable from the technology used to facilitate it; in many cases, to paraphrase Marshall McLuhan, the medium is the abuse.⁴⁵ Katsh and Rabinovich-Einy’s online dispute resolution approach places the most vital questions about technology beyond dispute: Who reaps its benefits? What values are embedded in its design? Whose labor does it exploit? Should it exist at all? The technological-determinist framework shifts attention away from political and cultural questions of inequality and power toward questions of data and efficiency. The justice of technological determinism is, therefore, an impoverished form of justice. Justice cannot be achieved in the most destructive and widespread anti-social conflicts without an unflinching assessment of consent, power, and labor; reflection on the compromised nature of online intermediaries; and repudiation of technological determinism.

Part II of this Review examines how and why online dispute resolution can be productive for pro-social disputes. Part III argues that the technological determinism inherent in Katsh and Rabinovich-Einy’s vision of ODR makes its application to anti-social disputes counterproductive and destructive.

⁴⁴ See generally *It’s Not a Bug, It’s a Feature*, URBAN DICTIONARY (Sept. 14, 2012), <https://www.urbandictionary.com/define.php?term=It%27s%20not%20a%20bug%2C%20it%27s%20a%20feature> [<https://perma.cc/56HF-P452>].

⁴⁵ See MARSHALL MCLUHAN, UNDERSTANDING MEDIA 7 (1964).

II. PRO-SOCIAL DISPUTES

The purpose of *Digital Justice*, according to the authors, is “to clarify not only how technology generates disputes but how technology can be employed to resolve and prevent disputes” (p. 3). Increasingly sophisticated technology produces increasingly sophisticated conflicts, but our tools for addressing these conflicts have not evolved as quickly. The authors aim to address the gap between progressively complex disputes and stagnating dispute resolution systems (p. 3). The tool they propose for doing so is Online Dispute Resolution, or ODR (p. 178). Earlier work by Katsh and Professor Janet Rifkin explained that all successful dispute resolution systems can be conceptualized as a triangle whose sides represent three essential elements: convenience, expertise, and trust (p. 37).⁴⁶ ODR adherents emphasize the role that technology can play in enhancing each of these elements, serving as a “Fourth Party” to aid the human parties navigating dispute resolutions (p. 37). According to Katsh and Rabinovich-Einy, technology’s ability to enhance the expertise side of the triangle in particular will most likely be the key to encouraging increased use of ODR (p. 38).

Katsh and Rabinovich-Einy’s ODR perspective has much to recommend it. It provides a refreshing counterpoint to the laissez-faire, anti-regulatory framework of many internet scholars and activists. The authors’ detailed knowledge of internet history (pp. 25–38), in particular of its regulation by the government (pp. 15–17, 26), keeps them from falling prey to the faux-libertarianism of many Silicon Valley enthusiasts.⁴⁷ For example, when they discuss removals of Google search engine results due to the much-maligned European “Right to be Forgotten” rulings, the authors point out that accidentally disclosed data from Google’s Transparency Report showed that “unlike the selective requests disclosed by Google relating to criminals and public figures, the vast majority of requests came from the general public relating to more mundane privacy-related concerns” (p. 117). Katsh and Rabinovich-Einy take for granted what many internet enthusiasts are unable or unwilling to recognize: that there is no progress, technological or otherwise, without regulation, and that quietism is no answer to the increasingly complex conflicts generated by evolving technology. Rather, they argue that “the faster these new problems have grown, the more urgently we need to prioritize as a society thinking about how to prevent and resolve them” (p. 15). Their project is energetically interventionist, seeking ways to maximize the benefits of innovative technology while minimizing its

⁴⁶ The authors cite ETHAN KATSH & JANET RIFKIN, *ONLINE DISPUTE RESOLUTION* 73–92 (1st ed. 2001).

⁴⁷ See Mary Anne Franks, *Unwilling Avatars: Idealism and Discrimination in Cyberspace*, 20 COLUM. J. GENDER & L. 224, 234–37 (2011).

costs. The authors' primary goal seems to be quite simple and unimpeachable: to help people achieve a happier and more productive relationship to the technology that increasingly dominates their lives (p. 3).

According to Katsh and Rabinovich-Einy, the early years of cyberspace, from 1969 to 1992, were relatively harmonious. This was likely due to the fact that the small fraction of the population accessing the internet during that time was mostly made up of academics or military researchers; commercial activity on the internet was banned; and the World Wide Web, invented in 1989, was completely text-based (pp. 26, 29). Things began to change in 1992, when the ban on commercial activity was lifted, and, shortly after, the first web browsers capable of displaying images and the first internet service providers (ISPs) appeared (p. 29).

Once internet disputes started to emerge in earnest, it quickly became clear that new resolution processes were needed in order for online activity to thrive. The number of disputes that arise today in e-commerce alone is staggering: according to Katsh and Rabinovich-Einy, "[i]t has been estimated that disputes occur in 3–5 percent of online transactions, leading to over seven hundred million e-commerce disputes in 2015" (p. 67). One of the recurring themes of *Digital Justice* is how ill-equipped the court system is to deal with the exponential rise in disputes: "No one — neither the courts, nor alternative processes — is prepared to handle the volume, variety, and character of disputes that are a by-product of the levels of creative and commercial activity happening online today" (p. 14). Exploring and developing the potential for technology to provide alternative, innovative routes to justice, they argue, could well "transform[] our very understanding of the meaning of *justice*" (p. 165).

The authors' online dispute resolution approach emphasizes the virtues of convenience (p. 37), speed (pp. 74–75), user control (pp. 120–21), preferences for systematic (pp. 34–36) and preventive (pp. 17–20) efforts over discrete and responsive ones, and, of course, technology's capacity to enhance all of these virtues (pp. 37–38).

One of the greatest benefits of ODR, according to Katsh and Rabinovich-Einy, is its ability to increase access and convenience. ODR "allow[s] communication at a distance, and asynchronously — with participation at any time," which "remove[s] many long-established physical constraints or boundaries of time and space" (p. 37). In other words, ODR makes it possible for individuals to avoid the cost and inconvenience of travel and provides scheduling flexibility (p. 37). This in turn allows resolution processes to proceed more quickly. Speed is critical to effective dispute resolution, Katsh and Rabinovich-Einy explain, citing an eBay study that revealed that "buyers preferred to lose their case quickly rather than have the resolution process go on for an extended

period of time” (pp. 74–75).⁴⁸ ODR also broadens the control users have over the dispute resolution process. ODR tools help “create a space for users to discuss problems, feelings, and desired outcomes,” “allow[] . . . for direct user-to-user negotiation,” and help the “process [be] tailored to the characteristics of the parties and the dispute” (p. 121).

Finally, ODR’s directive of “using technology to anticipate categories of disputes and design preventive systems” encourages the development of system-wide procedures over singular tools and preemptive tactics over after-the-fact responses (p. 19). Katsh and Rabinovich-Einy describe the importance of gathering data “that reveals patterns of disputes and provides opportunities to both facilitate and monitor consensual agreements, thus making disputes in the future less likely” (p. 35).

These are features well suited to pro-social disputes involving people who have a common interest, whether that is commerce, amicable separation, orderly work relations, or the maintenance of accurate records. One of the key characteristics of pro-social disputes is that all parties are generally invested in resolving the dispute quickly and efficiently. When the interests of powerful corporate or institutional entities roughly align with those of individual parties — the online bookseller wants customers to return, the courthouse operates more efficiently with accurate records, the hotel has a strong incentive to avoid bad reviews — there is a strong foundation for effective and just dispute resolution.

Katsh and Rabinovich-Einy’s conception of disputes is expansive. In *Digital Justice*, the authors bring the perspective of ODR to bear on a variety of topics — e-commerce, health care, the labor economy, courts and public institutions, and social media — treating subjects as diverse as medical record errors, online shopping refunds, abusive language in video games, and revenge porn as disputes. While their ODR approach holds much promise for pro-social disputes, it is ill-suited to social media disputes that are deeply antagonistic and one-sided.

III. ANTI-SOCIAL DISPUTES

The features of the online dispute resolution approach presented in *Digital Justice* that make it suitable for pro-social disputes often produce perverse results in anti-social disputes. This is because the underlying conditions of many anti-social disputes are fundamentally different from those of pro-social disputes. They are one-sided, antagonistic, and involve dramatic disparities of power as well as unjustifiable allocations of burdens and benefits. They are better characterized as attacks than disputes. The most damaging and widespread social media conflicts, including horrific Facebook Live videos, revenge porn, online harassment campaigns, violent propaganda, conspiracy theories, and “fake

⁴⁸ The authors cite AMY J. SCHMITZ & COLIN RULE, *THE NEW HANDSHAKE: ONLINE DISPUTE RESOLUTION AND THE FUTURE OF CONSUMER PROTECTION* 55 (2017).

news,” almost always involve involuntary interactions. Unlike buyers, sellers, and middlemen coming together online to engage in commerce, divorcing parents attempting to work out a child custody arrangement, or medical health professionals endeavoring to provide the best treatment for their patients, anti-social conflicts involve murderers seeking audiences for their violence, revenge porn site owners profiting from sexual humiliation, and conspiracy theorists hounding the parents of dead children.

In anti-social disputes, there is no presumption that all the parties are invested in resolving the dispute quickly. In fact, one party may be deeply committed to dragging out the conflict as long as possible, the better to inflict prolonged harm on the other party. In such cases, ODR’s features of speed and convenience will be irrelevant or ineffective. The emphasis on user control can be used to perverse effect in anti-social disputes, allowing abusers to flood reporting systems with disingenuous complaints and false information. As Katsh and Rabinovich-Einy themselves note, a “new ODR system can also be its own source of disputes, when the reporting system is abused by revenge-driven users reporting on other users” (p. 121).

The powerful corporations that provide the technology and the platforms for these attacks often have few incentives to stop them, and in some cases are incentivized to ignore or aggravate them. Facebook, Google, Twitter, and others are not like hospitals or courthouses or libraries; they are not even like Amazon or eBay. They produce nothing and sell nothing except advertisements and information about users, and conflict among those users may well be good for business. As documentary filmmaker and activist Astra Taylor writes in *The People’s Platform*, they are “commercial enterprises designed to maximize revenue, not defend political expression, preserve our collective heritage, or facilitate creativity, and the people who work there are private employees, not public servants.”⁴⁹

A. *Technological Determinism*

When, as described above, Katsh and Rabinovich-Einy offer qualified praise for Facebook for attempting to meet challenges posed by “anti-social media” (p. 109), they erase Facebook’s outsized role in creating those very challenges. More troublingly, the authors of *Digital Justice* speak approvingly of the fact that Facebook created its dispute resolution system using “compassion research” connected to the highly controversial “Facebook Experiment” (p. 120). In that experiment, Facebook manipulated the news feeds of approximately 700,000 users

⁴⁹ ASTRA TAYLOR, *THE PEOPLE’S PLATFORM: TAKING BACK POWER AND CULTURE IN THE DIGITAL AGE* 221 (2014).

without their consent to study the impact of positive and negative feeds on their moods.⁵⁰ Katsh and Rabinovich-Einy write:

While the Facebook experiment received harsh criticism in light of lack of informed consent to the experiment, the compassion research project resulted in the launching of a dispute resolution system through which users can alert their friends to content that is offensive to them, explaining how it makes them feel and what actions they would like to be taken by their counterpart. (p. 120)

In other words, the authors seem to believe that the fact that the experiment produced useful results — at least to Facebook — outweighs the fact that it obtained this information by engaging in nonconsensual manipulation of users' emotions.

New York Times writer Farhad Manjoo expressed a similar sentiment in 2014, arguing that there was an upside to Facebook's manipulation:

It is only by understanding the power of social media that we can begin to defend against its worst potential abuses. Facebook's latest study proved it can influence people's emotional states; aren't you glad you know that? Critics who have long argued that Facebook is too powerful and that it needs to be regulated or monitored can now point to Facebook's own study as evidence.⁵¹

While Manjoo acknowledged that it was "problematic" that Facebook users did not consent to being experimented on, he asked, "if every study showing Facebook's power is greeted with an outcry over its power, Facebook and other sites won't disclose any research into how they work. And isn't it better to know their strength, and try to defend against it, than to never find out at all?"⁵²

It is worth noting that we know of this study only because Facebook voluntarily chose to make it public — in other words, we know only what Facebook wants us to know. Manjoo sees only two choices: for the public to be told about being used as unwitting guinea pigs for Facebook or for the public to be permanently kept in the dark about what Facebook is doing with us. The obvious third option — that Facebook should not experiment on its users without consent — is not even on the table. Like Katsh and Rabinovich-Einy, Manjoo appears to think that we should both believe and be appeased by the fact that Facebook will use its illicitly obtained information to advance users' interests in some way.

Even as Katsh and Rabinovich-Einy acknowledge that Facebook's dispute resolution system could be considered a form of "unpaid outsourcing, rendering users themselves responsible for 'cleaning up' the internet" (p. 120), they suggest that this is the best one can expect when

⁵⁰ Farhad Manjoo, *A Bright Side to Facebook's Experiments on Its Users*, N.Y. TIMES (July 2, 2014), <https://www.nytimes.com/2014/07/03/technology/personaltech/the-bright-side-of-facebook-social-experiments-on-users.html> [https://perma.cc/4EJV-3TBW].

⁵¹ *Id.*

⁵² *Id.*

“platforms either adopt a ‘hands-off’ approach or fail to deal with the problem effectively” (p. 120). The question they leave unasked is why these are the only possible choices. Why can we not expect Facebook to take it upon itself to find an effective, direct way to handle the problem it helps create?

Facebook’s response to the public outcry over the murders and rapes streamed on Facebook Live is a prime example of Facebook *both* taking a “hands-off” approach *and* failing to provide an effective response. As discussed above, adding 3000 moderators to help review content flagged by users might reduce the platform’s response time, but it will not solve the problem.⁵³ As writer Steve Coll observed in the *New Yorker*, “better software and detection tools might prevent some broadcasted suicides or violence but they cannot possibly stop all of it.”⁵⁴ Facebook’s response suggests that it either has no idea how to handle the problem it created, does not care, or both. If we care about conflict prevention, as Katsh and Rabinovich-Einy suggest we should, Facebook Live should be a textbook case of how not to behave. Facebook has helped unleash the scourge of live-streamed murders, rapes, and assaults on the world without taking any real responsibility for doing so.

And indeed, why would it? There is little incentive for largely unregulated, immensely profitable corporations to keep dangerous but lucrative products out of the public sphere. There seems to be little political or public will to demand that they do so. When a new technological product is made available to the public, the public quickly becomes attached to it, no matter how much harm it causes. The possibility of removing it from the market is quickly rejected or dismissed outright. “The best way to prevent a graphic video from being seen is to never let it be uploaded in the first place,” writes reporter Emily Dreyfuss in *Wired*.⁵⁵ But, she immediately concludes, “if you had to wait for Facebook’s approval of your video of a cat on a vacuum, you’d just post that video somewhere else. Facebook would alienate a large constituency of people who want the ability to immediately and easily share their lives. And Facebook can’t afford that.”⁵⁶

⁵³ Kantrowitz, *supra* note 12 (“As long as Facebook maintains a truly live product, it probably can’t prevent violence from airing in its feeds. Indeed, the violent videos already broadcast on Facebook Live make clear the exceedingly difficult challenge the company faces in managing them. Many start out calmly enough only to abruptly erupt into gunfire or violence.”).

⁵⁴ Steve Coll, *Facebook and the Murderer*, *NEW YORKER* (Apr. 18, 2017), <https://www.newyorker.com/news/daily-comment/facebook-and-the-murderer> [https://perma.cc/WR7Z-RLLB].

⁵⁵ Emily Dreyfuss, *AI Isn’t Smart Enough (Yet) to Spot Horrific Facebook Videos*, *WIRED* (Apr. 18, 2017, 9:39 PM), <https://www.wired.com/2017/04/ai-isnt-smart-enough-yet-spot-horrific-facebook-videos/> [https://perma.cc/NN79-SZKP].

⁵⁶ *Id.*

In fact, as one of the wealthiest companies in the world, Facebook can afford pretty much anything.⁵⁷ Facebook could certainly afford to take the time to think about how to address harmful uses of its products before making them available to the public. According to several sources, Facebook rushed into the live-streaming-video business and was caught off guard by the problems it created.⁵⁸ Zuckerberg pulled around 100 employees from other projects in early 2016 and instructed them “to work around the clock to roll out Facebook Live” in two months.⁵⁹ As the *Wall Street Journal* reported, “[a]t traditional companies, major product launches often take years. Technology firms, and Facebook in particular, emphasize speed even though they know it means there will be problems to iron out later,”⁶⁰ living up to Facebook’s one-time motto, “Move fast and break things”⁶¹ (and its silent corollary, “Make Lots of Money”). According to one source within the industry:

In the desire to push Live out to as many people as possible, there were a lot of corners that were cut. And when you take a fail-fast approach to something like live-streaming video, it’s not surprising that you come across these scenarios in which you have these huge ethical dilemmas of streaming a murder, sexual violence or something else.⁶²

Facebook could certainly afford to pass on a product if it determines that it cannot be used safely. There is no reason to assume that Facebook’s continued existence and profitability depend upon the rollout of untested and potentially hazardous new features. While it is no mystery why Facebook and other largely unregulated corporations would recklessly embrace potential revenue streams regardless of the harm they might produce, that is no reason for the general public to simply defer to this desire.

Conflicts inspired by Facebook Live demonstrate the limits of Katsh and Rabinovich-Einy’s ODR framework in situations involving extreme disparities of power and a lack of shared interests. The dispute resolution system Facebook created for young real-life friends to articulate hurt feelings to one another (p. 120) may indeed be admirable in limited contexts, but it is useless to the victims of aggressively anti-social attacks from abusive spouses, stalkers, white supremacists, or performative criminals. Katsh and Rabinovich-Einy’s framework assumes that technology, no matter how harmful, poorly designed, or hastily introduced, should exist, and all that can be done in the wake of such technology is

⁵⁷ See, e.g., Matt Egan, *Facebook and Amazon Hit \$500 Billion Milestone*, CNN (July 27, 2017, 10:29 AM), <http://money.cnn.com/2017/07/27/investing/facebook-amazon-500-billion-bezos-zuckerberg/index.html> [https://perma.cc/YQR2-KUTA].

⁵⁸ Seetharaman, *supra* note 15.

⁵⁹ *Id.*

⁶⁰ *Id.*

⁶¹ *Id.*

⁶² Alvarez, *supra* note 12.

to attempt to mitigate its worst effects. The ODR framework encourages the use of technology to resolve the conflicts technology creates — the “Fourth Party” doctrine of ODR (p. 37) — but not to reject the technology itself or to even seriously question its source. While this approach may serve the interests of tech companies, it does not serve the interests of justice.

Once we refuse to accept the inevitability or the supremacy of technology, we can think more broadly about how to address the conflicts it creates. For example, there are several ways that Facebook could try to respond substantively to the criticism of Facebook Live. The company could implement a delay similar to that commonly used in live television broadcasts.⁶³ Or Facebook could impose some limitations on who can use the Facebook Live feature (the feature had in fact originally been restricted to celebrities).⁶⁴ For instance, it could restrict use of the feature to users who have undergone an interactive training session that highlights permissible and impermissible uses of the feature. It could revoke use of the feature from users who engage in harassing or other abusive behavior.

For broader-based solutions, we would need to move away from internet exceptionalism and begin holding companies accountable for the harmful behavior their platforms facilitate.⁶⁵ Companies like Facebook could be compelled to provide some form of cost-benefit analysis to the public or to a regulatory agency to explain why the rewards of a given feature outweigh the risks. If they fail to do so, these companies could be prevented from making the feature available to the public. As Coll writes:

It is true that the advent of social media cannot be undone, any more than television could be regulated in a way that would fully prevent terrorists from exploiting it. Yet every corporation is vulnerable — maybe a better word is accountable — when the choices it makes harm others, particularly when the harm occurs in pursuit of profit.⁶⁶

B. Hidden Labor

The technological-determinist framework tends to erase or obscure the question of who should be responsible for dealing with the negative consequences of technology. Facebook’s user conflict resolution process, which Katsh and Rabinovich-Einy find so worthy of praise, shifts the burden of conflict resolution from the company to users. But asking users to provide more free labor to solve the problems that the company helped create is not justice, but exploitation.

⁶³ Alvarez, *supra* note 12.

⁶⁴ Seetharaman, *supra* note 15.

⁶⁵ For a specific example of such an approach, see Mary Anne Franks, *Sexual Harassment 2.0*, 71 MD. L. REV. 655, 677–704 (2012).

⁶⁶ Coll, *supra* note 54.

That kind of exploitation is central to Facebook's vision. Mark Zuckerberg's lengthy February 2017 manifesto about the company's future touted the reliance on individual user labor as liberating and democratic:

The idea is to give everyone in the community options for how they would like to set the content policy for themselves. Where is your line on nudity? On violence? On graphic content? On profanity? What you decide will be your personal settings. We will periodically ask you these questions to increase participation and so you don't need to dig around to find them. For those who don't make a decision, the default will be whatever the majority of people in your region selected, like a referendum. Of course you will always be free to update your personal settings anytime.

With a broader range of controls, content will only be taken down if it is more objectionable than the most permissive options allow. Within that range, content should simply not be shown to anyone whose personal controls suggest they would not want to see it, or at least they should see a warning first. Although we will still block content based on standards and local laws, our hope is that this system of personal controls and democratic referenda should minimize restrictions on what we can share.⁶⁷

Facebook's strategy of presenting exploitation as liberation is remarkably successful. In *Move Fast and Break Things*, film producer Jonathan Taplin writes:

Mark Zuckerberg's greatest insight was that the human desire to be "liked" was so strong that Facebook's users would create all the content on the site for free. In 2014, Facebook's 1.23 billion regular users logged in to the site for seventeen minutes each day . . . that's more than 39,757 years of time collectively spent on Facebook in a single day. That's almost fifteen million years of free labor per year. Karl Marx would have been totally mystified.⁶⁸

Astra Taylor writes that "society's increasing dependence on free labor" is not only immoral but also "glosses over the question of who benefits from our uncompensated participation online."⁶⁹ And benefit Facebook does: the company's first-quarter earnings in 2017 were "\$10.8 billion in revenue, up 49 per cent on the same quarter last year, which gave it a profit of \$4 billion for the first quarter of the year."⁷⁰ Facebook does not share those massive benefits with the users that create the content that drives its platform, to say nothing of attempting to compensate the people whose lives are ruined by the abuse Facebook helps facilitate.

⁶⁷ Mark Zuckerberg, *Building Global Community*, FACEBOOK (Feb. 16, 2017), <https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10103508221158471/> [<https://perma.cc/XRZ3-8X4X>].

⁶⁸ JONATHAN TAPLIN, *MOVE FAST AND BREAK THINGS: HOW FACEBOOK, GOOGLE, AND AMAZON CORNERED CULTURE AND UNDERMINED DEMOCRACY* 150 (2017).

⁶⁹ TAYLOR, *supra* note 49, at 65.

⁷⁰ Chester, *supra* note 29.

Katsh and Rabinovich-Einy write approvingly of the efforts of companies such as Facebook and Twitter to improve their responses to online conflicts (p. 122). But many of the improvements that these and other major social media companies have made are the product of another kind of free labor, in the form of uncompensated work done by victims and advocates. For example, Facebook finally acknowledged the problem of violently misogynist content on the site and committed to making changes following an intense, months-long campaign led by Soraya Chemaly and the Everyday Sexism Project.⁷¹ The winning strategy for that campaign was to target the advertisers whose content appeared on graphic pro-rape and pro-domestic violence pages.⁷² Similarly, Facebook, Twitter, Microsoft, and Google banned nonconsensual pornography (a.k.a. “revenge porn”) from their sites and established policies to request the removal of private material only after years of research memos, presentations, meetings, and collaborations with non-profit organizations such as the Cyber Civil Rights Initiative.⁷³

In addition to free labor, Facebook, like most major tech companies, relies on low-paid labor for much of its content-moderation work.⁷⁴ As journalist Adrian Chen wrote in *Wired*, much content moderation is done in the Philippines, a country that “has maintained close cultural ties to the United States, which content-moderation companies say helps Filipinos determine what Americans find offensive. And moderators in the Philippines can be hired for a fraction of American wages.”⁷⁵ These companies appear to be largely unconcerned about the ethical implications of subjecting increasing numbers of employees to live videos of beheadings, child porn, rape, and torture as part of their daily job.⁷⁶ The long-term effects of exposure to this kind of content can be extremely damaging, and those performing this labor are often unprepared for the psychological toll it will take on them.⁷⁷

Facebook’s seeming lack of concern for the workers responsible for the vital, traumatizing work of content moderation was illustrated by a 2016 security lapse that exposed personal details of more than 1000

⁷¹ Laura Bates, *The Day the Everyday Sexism Project Won — And Facebook Changed Its Image*, THE INDEPENDENT (May 29, 2013, 6:40 PM), <https://www.independent.co.uk/voices/comment/the-day-the-everyday-sexism-project-won-and-facebook-changed-its-image-8636661.html> [<https://perma.cc/P82X-6SHU>].

⁷² *Id.*

⁷³ Mary Anne Franks, “Revenge Porn” Reform: A View from the Front Lines, 69 FLA. L. REV. 1251, 1270–72 (2017).

⁷⁴ Adrian Chen, *The Laborers Who Keep Dick Pics and Beheadings out of Your Facebook Feed*, WIRED (Oct. 23, 2014, 6:30 AM), <https://www.wired.com/2014/10/content-moderation/> [<https://perma.cc/N8HZ-UGRV>].

⁷⁵ *Id.*

⁷⁶ *Id.*

⁷⁷ *Id.*

Facebook content moderators who were tasked with screening out beheadings, child pornography, and terrorist propaganda. Of those affected, “around 40 worked in a counter-terrorism unit based at Facebook’s European headquarters in Dublin, Ireland. Six of those were assessed to be ‘high priority’ victims of the mistake after Facebook concluded their personal profiles were likely viewed by potential terrorists.”⁷⁸ One of these moderators described being treated as a “second-class citizen[.]” compared to Facebook employees.⁷⁹ According to the moderator, he received only two weeks of training before beginning his work, was paid fifteen dollars an hour, and was required to use his own personal Facebook account to log in to work. “‘They should have let us use fake profiles,’ he said, adding: ‘They never warned us that something like this could happen.’”⁸⁰

Facebook is now the sixth largest company in the world by market capitalization.⁸¹ It can more than afford to pay the workers who perform some of the platform’s most disturbing and important tasks better wages and to provide for their security and mental well-being. It can also, for that matter, do far more to absorb the costs generated by its various products, costs that are currently borne by vulnerable individuals and groups targeted for online abuse.

C. *Hidden Biases*

While Katsh and Rabinovich-Einy are impressed by some of the discrete tools developed by Facebook and Twitter, they are even more enthusiastic about the complex system of dispute resolution developed by Wikipedia, the largest free online encyclopedia (pp. 122–25). Wikipedia’s elaborate set of principles and rules for editing are indeed impressive. As the authors describe it, Wikipedia has developed:

[V]arious measures that would allow for constructive discussion and consensus-building, while ensuring that quality is maintained and abuse is addressed or prevented. These measures have included clear and predetermined rules governing the editing process; a multilayered ODR system; and a hierarchy of editors with varying levels of authority in determining editing disputes and regulating editor misconduct. (pp. 122–23)

Katsh and Rabinovich-Einy also find much to admire in Wikipedia’s conflict-prevention efforts. Wikipedia makes use of software programs

⁷⁸ Olivia Solon, *Revealed: Facebook Exposed Identities of Moderators to Suspected Terrorists*, THE GUARDIAN (June 16, 2017, 3:09 PM), <https://www.theguardian.com/technology/2017/jun/16/facebook-moderators-identity-exposed-terrorist-groups> [https://perma.cc/8J4U-UJDD].

⁷⁹ *Id.*

⁸⁰ *Id.*

⁸¹ *Here Are the 20 Largest Companies in the World by Market Cap*, BUS. TECH (July 12, 2017), <https://businesstech.co.za/news/business/184817/here-are-the-20-largest-companies-in-the-world-by-market-cap/> [https://perma.cc/Q9L8-22SZ].

called “bots” to assist in editing vandalism (p. 125). The automated processes make it possible for Wikipedia editors to quickly detect vandalism operations that a human editor might never catch (p. 126).

Wikipedia’s ODR system allows users with no experience in dispute resolution to create some of the site’s processes, a feature that the authors find particularly meaningful (p. 124). Katsh and Rabinovich-Einy write:

It is the combination of the diversity in dispute system designer identities as well as the unique features of digital technology that have rendered Wikipedia’s dispute resolution processes different from the traditional offline dispute resolution landscape. This is evidenced in the relaxed attitude toward confidentiality . . . as well as the ease with which input can be drawn from a wide range of editors. (p. 124)

Wikipedia’s far greater success in using ODR compared to other social media platforms makes sense given the nature of the site. Wikipedia editors are bound together by a common purpose, even if frequent disputes arise among them about how best to carry out that purpose. Wikipedia disputes, then, appear to be more pro-social rather than anti-social interactions. The authors’ admiration for Wikipedia’s moderation and dispute resolution system is echoed by others who study online communities, including Professor James Grimmelman: “Wikipedia takes its community democracy as seriously as it can.”⁸²

But appearances of diversity and democracy can be deceiving. A whopping eighty-four percent to ninety-one percent of Wikipedia editors are male.⁸³ One of the reasons for this gender disparity may be related to the “gender gap” in leisure time. In the United States as well as across the world, women have less free time to engage in unpaid, noncare activities than men do.⁸⁴ According to a 2013 Pew Research survey, men spend three hours more a week than women do on leisure activities.⁸⁵ But another reason is that the tiny number of women who do edit the site have experienced prolonged and severe sexual harassment by fellow

⁸² James Grimmelman, *The Virtues of Moderation*, 17 YALE J.L. & TECH. 42, 87 (2015).

⁸³ Jenny Kleeman, *The Wikipedia Wars: Does It Matter If Our Biggest Source of Knowledge Is Written by Men?*, NEW STATESMAN (May 26, 2015), <https://www.newstatesman.com/lifestyle/2015/05/wikipedia-has-colossal-problem-women-dont-edit-it> [<https://perma.cc/ABK7-BM3C>].

⁸⁴ *Balancing Paid Work, Unpaid Work and Leisure*, ORG. FOR ECON. CO-OPERATION & DEV. (July 3, 2014), <https://www.oecd.org/gender/data/balancingpaidworkunpaidworkandleisure.htm> [<https://perma.cc/BPL2-8YYL>].

⁸⁵ Kim Parker & Wendy Wang, *Modern Parenthood: Roles of Moms and Dads Converge as They Balance Work and Family*, PEW RES. CTR. (Mar. 14, 2013), <http://www.pewsocialtrends.org/2013/03/14/modern-parenthood-roles-of-moms-and-dads-converge-as-they-balance-work-and-family/> [<https://perma.cc/U8XV-DVGY>] (“Men spend more time than women in leisure activities The gender gap in leisure time is bigger among men and women who do not have children in the house (37 hours per week for men vs. 32 hours per week for women). Among parents with children under age 18, fathers spend, on average, 28 hours per week on leisure activities, while mothers spend 25 hours on leisure.”).

editors and have been actively prevented from engaging in efforts to remedy the situation.⁸⁶

In 2014, a female Wikipedia editor using the pseudonym Lightbreather was invited to join the Gender Gap Task Force, “a project by Wikipedia editors to examine why so few women participate on the site and why there’s a lack of coverage of notable women.”⁸⁷ Male editors would repeatedly insert themselves into the discussions, challenging the need for the project. Lightbreather quit the task force after a few days. In 2015, she proposed “a women-only space on Wikipedia for female editors to support each other and discuss the specific barriers they face online”; Wikipedia users objecting to the effort announced in the “oppose” section of the discussion page that they would “fight this to the death.”⁸⁸

At one point, Lightbreather discovered that a fellow editor was posting images on a pornographic website and alleging that they were of her.⁸⁹ This was not the first time Lightbreather had experienced harassment related to the site; previously, she had asked Wikipedia administrators for a space to discuss how to enforce Wikipedia’s policy on civility. On a page that was set up to discuss her request, a fellow Wikipedia editor wrote, “The easiest way to avoid being called a cunt is not to act like one.”⁹⁰ Lightbreather asked Wikipedia’s Arbitration Committee (ArbCom), “a panel of 15 elected users who have the final say on all arguments between editors,” to take up her case against the editor who posted fake pornography photos of her.⁹¹ The ArbCom refused on the basis that taking the case might “‘out’ the editor that had posted the pictures, or link his username to his real name.”⁹² At the same time, another editor had opened a case against Lightbreather, accusing her of having a “battleground mentality” based on her complaints about other editors. ArbCom’s ruling in that case was to ban Lightbreather from editing for at least one year.⁹³

Journalist Jenny Kleeman, writing in the *New Statesman* in 2015, observed that “[i]nstead of being the egalitarian ‘sum of all human knowledge’” that Wikipedia’s founder, Jimmy Wales, had hoped for, the English version of the online encyclopedia “is mostly the sum of male

⁸⁶ See Emma Paling, *Wikipedia’s Hostility to Women*, THE ATLANTIC (Oct. 21, 2015), <https://www.theatlantic.com/technology/archive/2015/10/how-wikipedia-is-hostile-to-women/411619/> [<https://perma.cc/9MMC-NM3F>].

⁸⁷ *Id.*

⁸⁸ *Id.*

⁸⁹ *Id.*

⁹⁰ *Id.*

⁹¹ *Id.*

⁹² *Id.*

⁹³ *Id.*

knowledge.”⁹⁴ One editor found that almost 4400 female scientists who met the Wikipedia notability standards did not have a Wikipedia entry.⁹⁵ In 2013, a reporter discovered that an editor had removed every female novelist from a list of American novelists and put them in a separate list titled “American women novelists.”⁹⁶ According to historian Gina Luria Walker, Wikipedia does not look so different from the first edition of the *Encyclopedia Britannica*, which was written between 1768 and 1771 by 150 men and no women.⁹⁷ The volume “featured 39 pages on curing disease in horses, and three words on woman: ‘female of man.’”⁹⁸ As former Wikimedia Foundation contractor Sarah Stierch put it, “[w]hen white men have been editing history since day one, they don’t see this as a problem.”⁹⁹

Wikipedia’s much-praised moderation and dispute resolution practices give the appearance of diversity and of serving the interests of procedural justice. But behind this appearance is the reality of bias against women. Discovering this requires looking beyond appearances — beyond procedure — into the platform’s substantive commitments and practices.

CONCLUSION

“Disputes,” write Katsh and Rabinovich-Einy, “are the collateral damage of innovation” (p. 5). While that may be true, it is also true that we do not all share the same risk of becoming collateral damage and that we do not share equally in the spoils of innovation. While the online dispute resolution approach championed in *Digital Justice* holds great promise for improving procedural justice in commercial and institutional transactions, its idealization of technology makes it unsuitable for resolving anti-social interactions facilitated by social media. The internet is currently overrun with violence, threats, revenge porn, propaganda, and conspiracy theories. These conflicts, which disproportionately burden women and racial, religious, and sexual minorities, are facilitated by powerful tech corporations with little incentive to pour their considerable resources into eliminating them. Resolving or preventing these disputes requires the rejection of technological determinism and engagement with the reality of consent, power, labor, and compromised intermediaries.

⁹⁴ Kleeman, *supra* note 83.

⁹⁵ Paling, *supra* note 86.

⁹⁶ *Id.*

⁹⁷ *Id.*

⁹⁸ *Id.*

⁹⁹ *Id.*

“[T]he internet is broken,” Evan Williams, co-founder of Twitter and co-creator of Blogger, told the *New York Times* in May 2017.¹⁰⁰ “‘I thought once everybody could speak freely and exchange information and ideas, the world is automatically going to be a better place,’ Mr. Williams [said]. ‘I was wrong about that.’”¹⁰¹ Williams certainly wasn’t alone in the belief that technology was going to make the world a freer, more informed, and more interesting place. But he is one of the few now willing to admit that the fantasy of technological utopia is just that: a fantasy. Technology will not save us from ignorance, bigotry, greed, or violence. It is not technology that can define justice; it is justice that must define technology.

¹⁰⁰ David Streitfeld, “*The Internet Is Broken*”: @ev Is Trying to Salvage It, N.Y. TIMES (May 20, 2017), <https://www.nytimes.com/2017/05/20/technology/evan-williams-medium-twitter-internet.html> [<https://perma.cc/SV3R-HMMN>].

¹⁰¹ *Id.*