Washington and Lee Law Review

Volume 77 | Issue 4 Article 5

Fall 2020

The Digital Samaritans

Eldar Haber
Faculty of Law, Haifa University, ehaber@univ.haifa.ac.il

The Digital Samaritans

Eldar Haber*

Abstract

Bystanderism is becoming largely digital. If being subjected to perilous situations was once reserved almost solely for the physical world, individuals now might witness those in peril digitally from afar via online livestreams. New technological developments in the field of artificial intelligence (AI) might also expand bystanderism to new fields, whereby machines—not just humans—are gradually positioned to better compute their surroundings, thus potentially being capable of reaching a high statistical probability that a perilous situation is currently taking place in their vicinity. This current and future expansion of bystanderism into the digital world forms a rather new type of digital bystander that might challenge the legal and social meaning of bad Samaritan laws—legal duties to act on the behalf of others in a perilous situation by reporting the events or aiding those in the perilous situation, when the burden or risk of such aid is low. With the rise in the availability of livestreaming crimes on social media platforms, and the rise in AI capabilities, the current legal framework that governs bad Samaritans might become inappropriate in regulating social behavior and personal safety, which in turn might shift to the almost sole prerogative of

* Senior Lecturer, Faculty of law, University of Haifa; Faculty member, Center for Cyber, Law and Policy (CCLP) and Haifa Center for Law and Technology (HCLT), Faculty of law, University of Haifa. I am much grateful to Niva Elkin-Koren for fruitful discussion on this topic and to Gabriel Focshaner for his excellent assistance in research.

platform governance—transforming online users and platforms into becoming the new digital Samaritans.

Table of Contents

I. Introduction 1560

II. Good and Bad Samaritan Laws 1563

III. Online and Artificial Bystanders 1577
A. Online Bystanderism 1578
1. Online Communicators 1582
2. Online Viewers 1590
B. Artificial Bystanderism 1605

IV. Samaritans and the Rise of Platform Governance 1626

V. Conclusion 1644

I. Introduction

Bystander intervention is not rooted within common law jurisprudence.[1] With some exceptions, individuals in society are not generally obliged to aid or rescue others in perilous situations, even if the burden or risk of such aid is perceived to be low, as common law does not impose liability for nonfeasance.[2] Triggered by the public's moral outcry over media-reported events, whereby eyewitnesses of horrific crimes did not report or aid the victims, state legislators responded to bystander intervention in the form of bad Samaritan laws.[3] These laws were crafted generally to impose affirmative requirements on individuals to assist others in perilous

1. *See* Jay Silver, *The Duty to Rescue: A Reexamination and Proposal*, 26 WM. & MARY L. REV. 423, 424 (1985) ("With limited exceptions, there is no duty under Anglo-American law to lend personal assistance or to obtain help for persons in distress, or to warn of imminent danger.").

2. *See id.* (explaining that under U.S. law even an "expert swimmer, with a boat and a rope in hand" need not rescue a person who he sees drowning in front of him).

3. See *infra* note 34 and accompanying text for a discussion of public outcry prompting legislatures to adopt bad Samaritan laws.

situations or notify enforcement agencies, as long as the reporter or rescuer lacks any associated risks with such duty.[4]

Recent technological developments, however, might alter the ways individuals are subjected to criminal activities or other perilous situations in real time. As many social media platforms have enabled users to livestream almost any aspect of their lives, it was only a matter of time until numerous individuals took advantage of this technological feature to communicate violent criminal conduct, such as rape, assault, murder, and mass massacres, to name but a few examples.[5] If witnessing crimes was once solely confined to the kinetic realm, requiring physical presence of the spectator, the emergence of these *livestreaming crimes* challenges this notion. And this might only be the beginning, as bystanderism might soon also extend into the realm of machines as well.[6] With advancements in the field of artificial intelligence (AI), potential new forms of Samaritans may emerge as these AI-based devices might reach computational abilities to detect perilous situations in real time, and aside from gathering evidence for potential future proceedings, will be positioned to render aid from enforcement agencies directly.[7]

These current and potential technological developments create new categories of digital Samaritans—online and artificial ones.[8] Online bystanders are those individuals that either communicate or view perilous situations like livestreaming crimes—thus representing the current state of technology in this respect.[9] The second type of potential digital Samaritans, artificial bystanders, refers to AI-based devices and services that use their sensors and advanced speech recognition

4. *See infra* Part II.

5. *See infra* Part II.A.

6. *See, e.g.*, Christine Hauser, *In Connecticut Murder Case, a Fitbit is a Silent Witness*, N.Y. TIMES (Apr. 27, 2017), https://perma.cc/BCB3-X8NW (noting that in this homicide case the victim's movements tracked on her Fitbit "may be the key to solving her murder").

7. *See infra* Part III.

8. *See infra* Part III.

9. *See infra* Parts III.A–B.

and machine learning algorithms to accurately assess perilous situations like criminal activities that involve victims.[10]

This new form of digital Samaritans raises a host of legal, social, and moral questions. In the context of Samaritan laws, they require reevaluating the current legal and moral duties and perceptions imposed on individuals in society, and AI services, to render assistance for those in distress or in perilous situations, by reporting it or aiding those in peril in any other way.[11] Should policymakers react to these new forms of subjection to perilous situations like livestreaming crimes or the potential ability of machines to safeguard individuals' personal safety? And if so, how? Can the current legal framework of bad Samaritan laws, in states where applicable, be applied to this new form of Samaritans? What are the benefits and drawbacks of imposing such affirmative duties on online users and platforms, especially within the context of civil rights and liberties? And what are the broader normative consequences of the state's current reluctance to interfere in regulating online platforms directly in this respect—thereby triggering a social-technological-legal debate on the proper scope and ramifications of platform governance, and its potential reconstruction of democratic values and individuals' civil rights and liberties?

This Article addresses these and other timely questions regarding the role of both individuals and platforms to serve as digital Samaritans, thereby aiding members of society in perilous situations, while further broadening the emerging scholarly discussion on platform governance in general. It proceeds as follows. Part II traces the historical roots of bad Samaritan laws, while further differentiating and evaluating different duties that the law currently imposes on bystanders in the kinetic world. Part III introduces and discusses the proclaimed movement towards digital Samaritans. This Part is further divided into two, differentiating between online and artificial Samaritans, while respectively evaluating the plausibility and desirability of imposing affirmative duties to report or assist in the digital era. Part IV zooms out to discuss

10. *See infra* Part III.

11. *See infra* Part IV.

the rise of platform governance in shaping the moral duties of society, and the potential roles these platforms might assume in public enforcement. It argues that while the law should play a limited role in the shaping of digital Samaritan duties, the rise of platform governance, especially in the realm of public enforcement, could highly endanger democratic values and must be better confined by both policymakers and users. The final Part summarizes the discussion and further warns against the unregulated shaping of human rights and liberties by online platforms.

II. Good and Bad Samaritan Laws

Common law does not generally impose any duties on individuals to act on behalf of others in peril.[12] Tracing its historical roots, common law rationale was to impose liability only for malfeasance but not for nonfeasance.[13] There was no general duty to assist or rescue others, albeit with a few exceptions in instances of special relationships between the person in peril and the bystander.[14] Accordingly, and within the context of criminal law, the initial federal criminal code did not impose duties on individuals for not reporting perilous

12. In the words of William L. Prosser, describing the lack of a legal duty under Anglo-American law: "The expert swimmer, with a boat and a rope at hand, who sees another drowning before his eyes, is not required to do anything at all about it, but may sit on the dock, smoke his cigarette, and watch the man drown." Silver, *supra* note 1, at 424 (citing W. PROSSER & W. KEETON, THE LAW OF TORTS § 56, at 375 (5th ed. 1984))*; see* WAYNE R. LAFAVE, CRIMINAL LAW 215 (3d ed. 2000) ("A moral duty to take affirmative action is not enough to impose a legal duty to do so."); Melody J. Stewart, *How Making the Failure to Assist Illegal Fails to Assist: An Observation of Expanding Criminal Omission Liability*, 25 AM. J. CRIM. L. 385, 387 (1998) ("[T]he common law rule regarding omission liability imposes no general legal duty or obligation upon one to act on behalf of anyone in peril.").

13. *See* Jessica R. Givelber, Note, *Imposing Duties on Witnesses to Child Sexual Abuse: A Futile Response to Bystander Indifference*, 67 FORDHAM L. REV. 3169, 3173 (1999) ("[A]bsent a special relationship, an individual has no affirmative legal duty to rescue another person in a perilous situation.").

14. *See* Ernest J. Weinrib, *The Case for a Duty to Rescue*, 90 YALE L.J. 247, 247 (1980) ("This general rule rests on the law's distinction between the infliction of harm and the failure to prevent it.").

situations, and more specifically, crimes.[15] With some exceptions,[16] crime control at that time was reserved solely for federal and state agencies—generally excluding other members of society.[17] In other words, while assistance to those in distress could stem from a religious view or moral codes,[18] it was not generally a legal duty in common law jurisdictions.[19]

Modern embodiments of these then-mostly religious and moral codes first emerged in medieval Europe, initially within limited instances like common disasters, only to be later expanded during the nineteenth century to criminal duties to

15. *See* Matthew R. Hall, Note, *An Emerging Duty to Report Criminal Conduct: Banks, Money Laundering, and the Suspicious Activity Report*, 84 KY. L.J. 643, 643 (1995) ("[W]hile a misprision of felony statute has appeared in the United States criminal code since 1790, the courts have interpreted that law to punish active concealment rather than passive non-reporting.").

16. Scholars argued that common law did impose a legal duty to prevent a felony in the mid-thirteenth century, among other related duties. *See* Givelber, *supra* note 13, at 3175–76.

17. *Cf.* Sandra Guerra Thompson, *The White-Collar Police Force: "Duty to Report" Statutes in Criminal Law Theory*, 11 WM. & MARY BILL RTS. J. 3, 8 (2002) ("[The] latest reactions to the child abuse scandals in the Catholic Church highlight the tendency legislators have shown to respond to such crises by requiring people in certain professions to report suspicions of criminality to the police.").

18. *Compare* Zachary D. Kaufman, *Protectors of Predators or Prey: Bystanders and Upstanders Amid Sexual Crimes*, 92 S. CAL. L. REV. 1317, 1335–36 (2019) (tracing the religious and moral codes within faiths such as Christianity, Islam, and Judaism; and discussing scholars like Jeremy Bentham Burke, Cicero, Dante, Martin Luther King, Jr., and Elie Wiesel), *with* Heather Benzmiller, Note, *The Cyber-Samaritans: Exploring Criminal Liability for the "Innocent" Bystanders of Cyberbullying*, 107 NW. U. L. REV. 927, 944 (2013) (tracing civilian duties to rescue and report within ancient Egypt, Plato's Laws, and Ancient German laws).

19. *See, e.g.*, Buch v. Amory Mfg. Co., 44 A. 809, 811 (N.H. 1898) ("The duty to protect against wrong is, generally speaking and excepting certain intimate relations in the nature of a trust, a moral obligation only, not recognized or enforced by law."); Osterlind v. Hill, 160 N.E. 301 (Mass. 1928) (holding that a boat rental business was not liable for not aiding its drowning); Natalie Perrin-Smith Vance, Comment, *My Brother's Keeper? The Criminalization of Nonfeasance: A Constitutional Analysis of Duty to Report Statutes*, 36 CAL. W. L. REV. 135, 138 (1999) ("Historically, there has never been a legal penalty for failing to come to the aid of others in the United States.").

report in some European nations.[20] These then-new legal codes sparked much discussion, and academics began to focus attention on the potential role that witnesses of perilous situations might and should play.[21] Subsequently, several policymakers in non-Continental legal systems began to follow suit by imposing limited duties to assist law enforcement agencies on those who hold certain professions, likely to obtain information about certain crimes or misconduct.[22] Under this rationale, physicians might be required to report gunshot and knife wounds or report child and elder abuse or neglect.[23] Even parking garage owners might be obliged to report bullet holes in cars.[24] In these instances, the state imposes duties to report potential criminal activity in specific circumstances when special relationships between individuals exist.[25]

---

20. *See* Benzmiller, *supra* note 18, at 944

> During the nineteenth century, many European nations began to enact criminal causes of action based upon violations of the duty to rescue. By the turn of the twentieth century, roughly half of the continental legal systems recognized a general duty to rescue in criminal law. And in the twentieth century, European civil law jurisdictions began to embrace broader formulations of the duty to rescue.

21. *See, e.g.*, James B. Ames, *Law and Morals,* 22 HARV. L. REV. 97, 111–13 (1908) ("One who fails to intervene to save another from impending death or great bodily harm, follows as a consequence of his inaction, shall be punished criminally."); Francis H. Bohlen, *The Moral Duty to Aid Others as a Basis of Tort Liability*, 56 U. PA. L. REV. 217, 220 (1908) ("The difference between non-feasance and misfeasance while quite fundamental, is much less obvious. The final physical injury to the plaintiff may be the same whether defendant's alleged misconduct is an act of violence or a failure to protect him from the violence of others.").

22. Health professionals can serve as a good example, as they might encounter gunshot wounds, injuries, or other health issues related to violent crimes. *See* Thompson, *supra* note 17*,* at 9–10.

23. *See* Hall, *supra* note 15, at 645–46 (citing CAL. PENAL CODE § 11166 (West 2020) as one statute imposing such requirements).

24. *See id.* at 646 (citing 31 R.I. GEN. LAWS ANN. § 31-26-12 (2020) as one statute imposing such an obligation).

25. These special relationships relate often to the establishment of a reasonable care duty, e.g., between students and schools; employees and employers; tenants and landlords; children and parents; or other related professional duties. *See* Daniel B. Yeager, *A Radical Community of Aid: A Rejoinder to Opponents of Affirmative Duties to Help Strangers*, 71 WASH. U.

Still, such duties were rather limited as opposed to many countries with civil law jurisdictions,[26] as common law solely imposed a limited duty to report obligation in special relationships of dependence;[27] in contractual agreements;[28] when the bystander placed someone in the perilous situation;[29] by specific statutes;[30] or in other limited circumstances.[31]

While bystander intervention was not, to begin with, acknowledged as a legal duty,[32] it has since expanded rapidly to be mandatory in many common law jurisdictions (at first only in special relationships), imposing direct liability on those who

---

L.Q. 1, 9–10 (1993) (listing and exemplifying special relationships that create a legal duty to assist, like parent-child); Jay Logan Rogers, Note, *Testing the Waters for an Arizona Duty-to-Rescue Law*, 56 ARIZ. L. REV. 897, 898 (2014) (noting that common law jurisprudence does not establish a duty to rescue imperiled strangers). There are also specific acts that impose duties to report in various contexts of special relationships, such as suspected child abuse. *See* 34 U.S.C. § 20341 (requiring certain professionals who learn of child abuse to report).

26. *See* Rogers, *supra* note 25, at 900–01 ("[C]ontinental Europe has taken a different approach, and several nations impose legal penalties on nonrescuers.").

27. *See* Silver, *supra* note 1, at 425–26 ("These 'special relationships' include parent to child, spouse to spouse, common carrier to passenger, innkeeper to guest, storekeeper to customer, host to social guest, employer to employee, teacher or school official to student, and jailer to inmate.").

28. These could include, *inter alia*, lifeguards who are contractually obliged to rescue people from drowning, as well as "[f]iremen, police, nurses, baby-sitters, and many others [who] enter into agreements that require them to render aid." *Id*. at 426.

29. *See* Perrin-Smith Vance, *supra* note 19, at 138 (noting that an exception to the common law nonfeasance rule has been made where the accused put the victim in peril).

30. *See* Benzmiller, *supra* note18, at 945 (observing that American courts have traditionally been resistant to impose affirmative duties on bystanders unless those duties arose by special relationship, by contract, or by statute). These statutes could include, *inter alia*, "hit and run" laws and regulations that require drivers involved in an automobile accident to give assistance to those injured, regardless of fault. *See* Silver, *supra* note 1, at 425.

31. *See* Silver, *supra* note 1, at 426 (reporting that "one who negligently injures or imperils another has a duty to render reasonable assistance" and that "one who volunteers aid is under a duty to exercise reasonable care").

32. *See* Damien Schiff, *Samaritans: Good, Bad and Ugly: A Comparative Law Analysis*, 11 ROGER WILLIAMS U. L. REV. 77, 83 (2005) ("At common law no such duty developed.").

were simply eyewitnesses or became aware of certain felonies and did not report the crime or assist the victims.[33] Perhaps also partially influenced by global movements, these laws were mostly triggered by specific events—often a crime committed while bystanders had the riskless ability to report or aid the victim and did not—which led to public outcry to amend the law.[34] These laws are now articulated as "Samaritan laws,"[35] regulating bystander intervention in perilous situations.[36]

---

33. *See id.* at 81–88 (summarizing the historical development of legal duties imposed on bystanders in American jurisprudence); Thompson, *supra* note 17, at 11 (explaining the duties modern reporting statutes impose on individuals who obtain knowledge of information relating to the commission of certain offenses).

34. Many Samaritan laws emerged from specific cases that engendered public outrage. Perhaps the most highly publicized example, also sparking the first legislative response, is that of the murder of Kitty Genovese in New York City, where media reports indicated that thirty-eight witnesses of the crime did not report it to the police or attempt to aid her in any other way. Martin Gansberg, *37 Who Saw Murder Didn't Call the Police*, N.Y. TIMES, Mar. 27, 1964, at 1. Some, however, argue that this incident was highly exaggerated or even close to a myth. *See, e.g.*, Bertrand Crettez & Regis Deloche, *On the Optimality of a Duty-to-Rescue Rule and the Cost of Wrongful Intervention*, 31 INT'L REV. L. & ECON. 263, 263 n.2 (2011). Notably, the murder of Kitty Genovese also led to a new field of research in the psychology of prosocial behavior. *See* PATRICIA WALLACE, THE PSYCHOLOGY OF THE INTERNET 191 (2d ed. 2016). Another example is that of the "Steubenville rape," where reports indicated that bystanders did not intervene while witnessing the sexual assault of a sixteen-year-old girl. *See, e.g.*, Maia Szalavitz, *What Bystanders Can Do to Stop Rape*, TIME (Jan. 11, 2013), https://perma.cc/A6SE-TVZK; Sarah L. Swan, *Bystander Interventions*, 2015 WIS. L. REV. 975, 984 n.37 (noting that the "Steubenville rape" was a high-profile case which brought the "bystander effect" to the forefront of public attention). The final example is that of the state of California, which introduced a bad Samaritan law as a response to a horrific event that occurred in the women's restroom of a Las Vegas casino: reports indicated that a seven-year-old girl was sexually assaulted and strangled to death in the restroom, while videos from security cameras showed that two potential observers did not come to her rescue or report this event. *See* Perrin-Smith Vance, *supra* note 19, at 135 (describing the factual details of the incident).

35. *See* Stewart, *supra* note 12, at 388 n.9 (discussing the origin of the reference to "Samaritans" in this context, drawn from a New Testament Bible parable in the book of Luke in which a beaten man was eventually assisted by a person who lived in or came from Samaria).

36. *See id.* at 388 ("In an attempt to encourage people . . . to assist others in peril while leaving undisturbed the common law no-duty rule, legislatures began to enact Good Samaritan laws.").

Literature divides Samaritan laws into two groups—often labeling the Samaritan as either being "good" or "bad."[37] Good Samaritan laws generally fall under the rubric of tort law, providing some form of civil immunity from liability to individuals who render aid to another person,[38] seemingly injured or in threat of such injury, as long as such aid is not proven to be highly negligent or conducted in a reckless manner.[39] Bad Samaritan laws—the focus of this Article—are different in nature.[40] Falling mainly under the rubric of criminal law and relying mainly on the moral duty of individuals to aid others in grave danger, bad Samaritan laws impose one of two legal duties: to report or to offer rescue in perilous situations—often cases of felony—while the burden of risk of

37. *See, e.g.*, Schiff, *supra* note 32, at 81–88 (discussing the history and development of Samaritan laws and differentiating between those that impose an affirmative duty to assist (bad) and those that protect individuals rendering assistance to others (good)).

38. In some cases, good Samaritan laws will only apply in specific circumstances. One example is that of good Samaritan laws that apply solely to medical aid: one famous case involved a woman who caused permanent damage to her friend's spinal cord while removing her from a car following an accident. *See* Van Horn v. Watson, 197 P.3d 164, 165 (Cal. 2008). The woman was held liable in this instance, as the applicable good Samaritan law protected only those granting medical aid. *See id.* at 171; *see also* Mueller v. McMillan Warner Ins., 714 N.W.2d 183, 186 (Wis. 2006) (affirming the Wisconsin Court of Appeals' reversal of summary judgment granted in favor of defendant parents because the injurious care they provided to the plaintiff was not "emergency care," and because the relevant good Samaritan statute did not immunize non-emergency care). For other judicial decisions in the context of good Samaritans, see Swan, *supra* note 34, at 1024–26.

39. *See* Patricia Grande Montana, *Watch or Report? Livestream or Help? Good Samaritan Laws Revisited: The Need to Create a Duty to Report*, 66 CLEV. ST. L. REV. 533, 537 (2018) (defining the typical characteristics of good Samaritan laws). For more on good Samaritan laws, see generally Eric A. Brandt, Comment, *Good Samaritan Laws—The Legal Placebo: A Current Analysis*, 17 AKRON L. REV. 303 (1983); Carl V. Nowlin, Note, *Don't Just Stand There, Help Me!: Broadening the Effect of Minnesota's Good Samaritan Immunity Through* Swenson v. Waseca Mutual Insurance Co., 30 WM. MITCHELL L. REV. 1001 (2004).

40. *See* Ken Levy, *Killing, Letting Die, and the Case for Mildly Punishing Bad Samaritanism*, 44 GA. L. REV. 607, 610 (2010) (describing bad Samaritan laws as those under which any individual who could have attempted to save an individual in danger or could have notified professional rescuers of her plight would have been criminally prosecuted).

such aid is low.[41] Thus, unlike good Samaritan laws, bad Samaritan laws incorporate a duty to directly assist others—whether by reporting to authorities or offering aid—when another person is in immediate danger, all depending on the jurisdiction in question and other various circumstances.[42] Notably, the duty to rescue is currently more limited in scope than the duty to report, existing in only a handful of U.S. states[43] and often exempting assistance requirements when it might jeopardize the rescuer.[44]

Bad Samaritan laws, also categorized as bystander intervention statutes, vary between legal jurisdictions[45] and within states.[46] Some statutes are more specific in nature,

41. *See* Levy, *supra* note 40, at 617–19 (listing rationales for bad Samaritan laws). For more on the moral duty to report crime when costs are minimal, see Arthur Ripstein, *Three Duties to Rescue: Moral, Civil, and Criminal*, 19 L. & PHIL. 751, 752–54 (2000).

42. *See* Kaufman, *supra* note 18, at 1325 (explaining bad Samaritan laws); *see also* Michael Davis, *How Much Punishment Does a Bad Samaritan Deserve?*, 15 L. & PHIL. 93, 93 (1996) (same).

43. *See, e.g.*, MINN. STAT. § 604A.01 (2020) (imposing a duty to provide reasonable assistance, including obtaining or attempting to obtain aid from law enforcement or medical personnel, on an individual who knows that another person is exposed to or has suffered grave physical harm); HAW. REV. STAT. § 663-1.6 (2019) (same); VT. STAT. ANN. tit. 12, § 519 (2019) (same); WIS. STAT. § 940.34(2)(a) (2020) (same); 11 R.I. GEN. LAWS ANN. § 11-56-1 (2020) (same).

44. Specifically, while some states impose an affirmative duty to assist only when a person's life, health, or safety is in peril, other states more narrowly impose such a duty only when there is danger to life or when specific crimes create the peril. *See* Benzmiller, *supra* note 18, at 951. In addition, the scope of such an affirmative duty could change considering the risks to the victim and to the would-be rescuer. *See id.*

45. *See* Note, *The Failure to Rescue: A Comparative Study,* 52 COLUM. L. REV. 631, 639–42 (1952) (describing some of the variations between French and Anglo-American legal jurisdictions); Stewart, *supra* note 12, at 397–404 (describing variation in judicial treatment of the duty-to-assist among states); Kaufman, *supra* note 18, at 1342–48 (listing domestic and global duty to report statutes); Levy, *supra* note 40, at 616 (listing European countries with bad Samaritan laws).

46. At least some form of a criminal duty to rescue appears now in most Latin American and European countries. *See* Benzmiller, *supra* note 18, at 945. In France, for instance, citizens have the obligation to report all crimes. *See* Edward A. Tomlinson, *The French Experience with Duty to Rescue: A*

dealing solely with certain felonies like sexual violence,[47] physical injuries,[48] or murder,[49] while others are more general, applying to any perilous situation or criminal felony, and even failure to report the discovery of a human body.[50] Some statutes impose a duty only with regard to specific cohorts, for example, vulnerable populations like children or the elderly.[51] While federal law does not currently include a general bad Samaritan law provision,[52] all states have some form of a good Samaritan

Dubious Case for Criminal Enforcement, 20 N.Y.L. SCH. J. INT'L & COMPAR. L. 451, 482 (2000) ("The [French] courts . . . have always insisted that the defendant respond to a known peril in an appropriate fashion."); *see also* Code pénal [C. pén.] [Penal Code] art. 434-1 (2020) (criminalizing failure to report); *see generally* Kaufman, *supra* note 18, at 1337–39 (differentiating various duty-to-report laws).

47. *See, e.g.*, FLA. STAT. § 794.027 (2020) ("Duty to report sexual battery; penalties.").

48. *See, e.g.*, MINN. STAT. § 604A.01(1) ("A person at the scene of an emergency who knows that another person is exposed to or has suffered grave physical harm shall . . . give reasonable assistance to the exposed person."); VT. STAT. ANN. tit. 12, § 519(a) ("A person who knows that another is exposed to grave physical harm shall . . . give reasonable assistance to the exposed person unless that assistance or care is being provided by others.").

49. *See* MASS. GEN. LAWS ch. 268, § 40 (2020) ("Whoever knows that another person is a victim of aggravated rape, rape, murder, manslaughter or armed robbery and is at the scene of said crime shall . . . report said crime to an appropriate law enforcement official as soon as reasonably practicable.").

50. *See* OHIO REV. CODE ANN. § 2921.22(C) (West 2020) ("No person who discovers the body or acquires the first knowledge of the death of a person shall fail to report the death . . . .").

51. *See, e.g.*, Givelber, *supra* note 13, at 3180–84 (stating that mandatory child abuse reporting statutes exist at least in some form in all fifty states); CAL. PENAL CODE § 152.3 (West 2020) ("Observation of offenses against children.").

52. *See* Kaufman, *supra* note 18, at 1344 ("[N]o general Bad Samaritan statute exists in U.S. federal law."). The federal criminal code, however, includes some duties which resemble bad Samaritan duties, but are more limited in scope. *See, e.g.*, 18 U.S.C. § 4 (defining the crime of misprision as applying only to an affirmative act of concealment, unlike non-reporting); 18 U.S.C. § 2258A (duty to report child pornography by electronic communication service providers and remote computing service providers); 31 U.S.C. § 5318(g) (reporting of suspicious activity by financial institutions); 42 U.S.C. § 1986 (providing an action for neglect to prevent conspiracy). For examples of proposed federal laws, see Perrin-Smith Vance, *supra* note 19, at 140–41.

law,[53] and more than half of U.S. states currently have some form of a bad Samaritan statute.[54]

There are several normative justifications behind bad Samaritan laws. To name a few examples,[55] some justifications are rather utilitarian in scope: that these laws could aid in minimizing needless deaths and injuries;[56] provide an outlet for the public's moral outrage on specific events;[57] enable society to formally echo the message that bad Samaritanism is morally wrong and that our shared citizenship forms a social responsibility for each other;[58] promote public safety;[59] and provide deterrence for non-compliance.[60] The underlying

53. *See* Montana, *supra* note 39, at 537 ("All states, including the District of Columbia, currently have some form of a Good Samaritan statute.").

54. *See* Kaufman, *supra* note 18, at 1345 ("Bad Samaritan laws—whether duties to rescue, to report, or either—that do apply to most or all witnesses exist in twenty-nine states and Puerto Rico."). Notably, regulation is not limited to the scope of affirmative duties to rescue or report. Some policymakers, like the New York Metropolitan Transportation Authority and the DHS, have also raised public awareness to suspicious activities under their "If You See Something, Say Something" campaign. *See About the Campaign*, DEP'T HOMELAND SEC., https://perma.cc/KK85-GS32 ("'If You See Something, Say Something®' is a national campaign that raises public awareness of the indicators of terrorism and terrorism-related crime, as well as the importance of reporting suspicious activity to state and local law enforcement.").

55. For a full taxonomy of the rationales behind duty-to-report laws from various perspectives, see David A. Hyman, *Rescue Without Law: An Empirical Perspective on the Duty to Rescue*, 84 TEX. L. REV. 653, 655–56 (2006).

56. *See* Levy, *supra* note 40, at 627.

57. *See id*. at 627–28.

58. *Id*. at 628–29; *see* Steven J. Heyman, *Foundations of the Duty to Rescue*, 47 VAND. L. REV. 673, 680 (1994) ("[I]n its fully developed form the duty to rescue is . . . a social duty—an obligation owed not only to the community itself but also to the other members of that community."); Shalini Bhargava Ray, *The Law of Rescue*, 108 CALIF. L. REV. 619, 631 (2020) (discussing the implications of viewing the duty to rescue as based on rights associated with shared citizenship).

59. Benzmiller, *supra* note 18, at 949.

60. *See* Sungyong Kang, *In Defense of the "Duty to Report" Crimes*, 86 UMKC L. REV. 361, 368 (2017) (arguing that such non-compliance does lead to harm). *But cf.* Eric Mack, *Bad Samaritanism and the Causation of Harm,* 9 PHIL. & PUB. AFFS. 230, 241–59 (1980) (critiquing the argument that "the Bad Samaritan's inaction causes the harm which his action would have prevented"); Richard Epstein, *A Theory of Strict Liability*, 2 J. LEGAL STUD.

assumption is that criminal law's deterrence and stigma will regulate the behavior of those who otherwise would not have aided, and will potentially achieve other goals of criminal law like retribution for the wrongdoer.[61] Deontological justifications for bad Samaritan laws might rely on the proclaimed *malum in se* nature of the conduct, i.e., that simply not reporting a perilous situation, and perhaps most importantly a crime (or in some instances not aiding those in peril), is morally wrong behavior and thus should be criminalized.[62] Corrective justice arguments might rely, *inter alia*, on the notion that the law should condemn and punish immoral conduct.[63] While highly debatable, economic analysis of imposing such affirmative duties could also support bad Samaritan laws on efficiency grounds.[64]

On the other hand, there is much controversy regarding this form of liberty-limiting principle. Libertarians might argue that bad Samaritan laws excessively restrict individuals'

151, 190–95 (1973) (arguing that bad Samaritan laws assign liability where causation—a crucial condition for liability—is missing).

61. *See* Kang, *supra* note 60, at 367 ("[A] duty to report crime leads to increased detection of wrongdoings, thus achieving retribution-rehabilitation along with general and specific deterrence of crime."); *see also* Joshua Dressler, *Some Brief Thoughts (Mostly Negative) About "Bad Samaritan" Laws*, 40 SANTA CLARA L. REV. 971, 980–81 (2000) (suggesting two types of retributivist views on bad Samaritan laws: culpability-retributivists—who would rely on moral culpability, and harm-retributivists—who would rely on punishing those that cause social harm).

62. *See* Kang, *supra* note 60, at 365–70 (describing the moral aspects of Samaritan laws).

63. *See* Hyman, *supra* note 55, at 655 ("Corrective justice scholars will argue that the law should enforce common moral intuitions.").

64. While Landes and Posner generally argued that a duty to rescue could be inefficient, *see* William M. Landes & Richard A. Posner, *Salvors, Finders, Good Samaritans, and Other Rescuers: An Economic Study of Law and Altruism*, 7 J. LEGAL STUD. 83, 85–88 (1978) (modeling the duty to rescue and concluding that it generates economic inefficiency), others have disagreed or argued that this statement depends on various factors that must be considered. *See* Richard L. Hasen, *The Efficient Duty to Rescue*, 15 INT'L REV. L. & ECON. 141, 142 (1995) (challenging the assumption of Landes' and Posner's model that the populations of rescuers and victims are entirely separate); Bhargava Ray, *supra* note 58, at 631 (summarizing the controverted scholarly points of view on this topic).

liberty, individualism, and autonomy.[65] They might further suggest that such forced altruism is inherently wrong—even when reporting or rescuing is rather easily performed and even if rendering aid will not place the Samaritan in danger or peril[66]—thus they have a right to refrain from reporting or rescuing.[67] A consequentialist view might suggest that imposing affirmative duties to report or rescue could be harmful to society, as it will render rescuers reluctant or reduce the quality of potential rescues.[68] Some argue that these duties could infringe on the victim's privacy rights, which could be crucial in many instances, like that of sexual offenses.[69] Other arguments rely on the feasibility of identifying the bystanders[70] and the enforceability of these laws.[71]

Pragmatically, even when such a duty is imposed, many people will not obey the law for various reasons, e.g., not being

65. *See* Swan, *supra* note 34, at 999 (arguing that the resilience in American law of the no duty to rescue norm is connected to "principles of autonomy, individualism, and privacy").

66. *See* Perrin-Smith Vance, *supra* note 19, at 139 (noting courts' refusal to abandon the common law no duty to rescue rule, even in cases of easy rescue, and summarizing competing views in the literature).

67. *See* Levy, *supra* note 40, at 657 (explaining the libertarian objection to bad Samaritan laws by way of abstract syllogism); Liam Murphy, *Beneficence, Law, and Liberty: The Case of Required Rescue*, 89 GEO. L.J. 605, 632 (2001) (arguing that the American legal tradition is distinctly libertarian); Bhargava Ray, *supra* note 58, at 630 ("Where an individual creates or contributes to harm, that individual has some obligation to cure that harm; otherwise, the individual has no obligation. Thus, the law should not punish the mere failure to act, even if such a rule seems heartless or immoral.").

68. *See* Marin Roger Scordato, *Understanding the Absence of a Duty to Reasonably Rescue in American Tort Law*, 82 TUL. L. REV. 1447, 1466–76 (2008) (discussing potential pitfalls of a coercive rule); Bhargava Ray, *supra* note 58, at 630 ("On [the consequentialist] view, imposing a duty to rescue will likely backfire and reduce the total incidence of rescue, as 'reluctant rescuers' will eschew risky activities in order to avoid potential liability.").

69. *See* Benzmiller, *supra* note 18, at 947 ("A duty to report a crime may unacceptably violate the victim's privacy, especially for victims of sexual offenses.").

70. *See* Perrin-Smith Vance, *supra* note 19, at 139 ("Questions arise about the feasibility of identifying the individuals who fail to aid.").

71. *See id.* at 147 (questioning the effectiveness of duty to report statutes, noting that some prosecutors oppose them, and suggesting potentially more effective methods of encouraging reporting that don't rely on criminal punishment for failure to do so).

aware of it, fear of retaliation, low probability of being identified and prosecuted for such lack of aid, fear of legal ramifications or from interacting with law enforcement agents, and that aiding undermines other competitive norms.[72] Some argue that legal duties in this context are effectively limited to begin with, as they might lead to uncertainty about when and how to act,[73] and their impact on the willingness to help is low at best.[74] Others will vote against bad Samaritan laws, as it is potentially unconstitutional to criminalize nonfeasance,[75] and infringes upon the First Amendment as compelled speech, along with also potentially violating Fifth Amendment rights against self-incrimination.[76] As a final example of bad Samaritan controversy,[77] others argue that while it might be important to

72. *See* Eugene Volokh, *Duties to Rescue and the Anticooperative Effects of Law*, 88 GEO. L.J. 105, 106 (1999) (stating the limited influence duty to rescue laws have on different types of Samaritans); *see also* Swan, *supra* note 34, at 1006–07, 1026–28 (citing Sharita Forrest, *Study Examines Role of School Culture in Promoting Bullying, Bystander Intervention*, ILL. NEWS BUREAU (Aug. 11, 2014, 9:00 AM), https://perma.cc/VYJ5-YVP2 (discussing how school culture creates environments that promote or discourage bystander intervention)).

73. *See* Perrin-Smith Vance, *supra* note 19, at 139 ("A duty to aid others may also lead to an uncertainty about when to act and create problems of strangers intervening in the affairs of others.").

74. *See* Hyman, *supra* note 55, at 688 ("Stated bluntly, the available data provides no indication that imposing a duty to rescue has any effect whatsoever on the impetus to perform a non-risky rescue."); Volokh, *supra* note 72, at 106 (stating that the law's coercive force will be low because witnesses know they are unlikely to be identified if they just stay quiet).

75. For a full taxonomy on the constitutionality of criminalizing nonfeasance, see Perrin-Smith Vance, *supra* note 19, at 136–55.

76. *See* U.S. CONST. amends. I, V. The First Amendment was interpreted to also include the protection on refraining from speaking, i.e., that the government is restricted from compelling speech. Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Bos., 515 U.S. 557, 573 (1995) (stating that an important manifestation of the principle of free speech is that one who chooses to speak can also decide "what not to say"); Perrin-Smith Vance, *supra* note 19, at 148 ("The fact that a person witnesses a violent act by chance, should not be the basis for violating protected rights.").

77. Imposing affirmative duties could lead to harmful consequences like harm to victims due to the harms or risks to rescuers and these duties might "generate only a small number of additional rescue efforts" due to the discount in the significance of a purely altruistic act. *See* Swan, *supra* note 34, at 999

find proper ways to regulate human conduct in such a manner that will make members of society more willing to help others, it is not the task of the criminal law to do so.[78]

(stating that individualism holds a privileged position in American society and as a result the legal norm of no duty to rescue remains supreme in the vast majority of states); Scordato, *supra* note 68, at 1464–69 (discussing how the practical benefits of an affirmative duty of reasonable rescue in tort law would be limited because noncompliers would be unlikely to change their behavior).

78. The question of when criminal law should extend to a specific conduct is complex and has been lengthily debated in academic literature for over 150 years. *See* JOHN STUART MILL, ON LIBERTY 8–9 (Gertrude Himmelfarb ed., 1974) (1859) (explaining liberty had a long lineage before Mill). Some theories have focused on the notion of potential or actual *harm* to individuals and to society, while other focused on the *wrongful nature* of the conduct. *See id.* at 8. John Mill is famously known for suggesting that the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. *See id.* at 8–9. As for Samaritan laws, one might argue that they are generally designed to aid in preventing, or at least reducing, harm to others, thus they could be held as fulfilling Mill's harm principle's requirements. *See id.*; Benzmiller, *supra* note 18, at 949. Others, however, argued that if the individual did not cause the actual harm, it might not fulfill Mill's requirements. *See, e.g.*, Robert Justin Lipkin, Comment, *Beyond Good Samaritans and Moral Monsters: An Individualistic Justification of the General Duty to Rescue*, 31 UCLA L. REV. 252, 257 (1983); Perrin-Smith Vance, *supra* note 19, at 139. Joel Feinberg has further developed Mill's harm principle, suggesting, *inter alia*, that a failure to perform a low-risk rescue is harmful for the victim, thus might fulfill the harm principle requirements. *See* JOEL FEINBERG, HARM TO OTHERS 136–39 (1984); Bhargava Ray, *supra* note 58, at 14. By now, the harm principle is often considered as merely one element to consider when criminalizing conduct. *See* FEINBERG *supra*, at 137–38 (providing a diagram that introduces other possible effects on one person's conduct on another person's interests). The principle was further developed by many scholars over time, and perhaps mostly by Feinberg, arguing that "[i]t is always a good reason in support of penal legislation that it would be effective in preventing (eliminating, reducing) harm to persons other than the actor (the one prohibited from acting) *and* there is no other means that is equally effective at no greater cost to other values." JOEL FEINBERG, HARMLESS WRONGDOING xix (1988) (emphasis in original). Other approaches to criminalization, like that of *legal moralism*, would advocate for the criminalization of a morally wrongful conduct, perpetrated with a culpable state of mind due to its wrongfulness. *See* ANTONY DUFF ET AL., TOWARDS A NORMATIVE THEORY OF THE CRIMINAL TRIAL 17 (2007). Here, one might argue that even without fulfilling the harm principle, not aiding those in grave danger, while they could easily and harmlessly do so, is an immoral behavior, thus it might be criminalized on these grounds. *See* Bhargava Ray, *supra* note 58, at 15. Another approach emphasizes the

Whether policymakers should enact bad Samaritan laws, or even more broadly categorized, affirmative altruistic duties, has been lengthily scrutinized and debated in academic literature.[79] Notably, there are both considerable benefits and drawbacks that policymakers must take into consideration before imposing any affirmative duties on individuals to partake in law enforcement, even when such participation is minimal and easy, at best, and does not carry actual risks or inflicts harm on the reporter.[80]

---

protection of social values, suggests that criminalizing behavior should mostly rely on identifying an important *protected social interest*, which could justify the reason why a particular conduct should be prohibited by criminal law. *See, e.g.*, Mordechai Kremnitzer & Khalid Ghanayim, *Proportionality and the Aggressor's Culpability in Self-Defense*, 39 TULSA L. REV. 875, 879 (2003). In recent years, structured principled approaches to criminalization were suggested by few scholars, including myself. Examining bad Samaritan laws under these approaches, however, extends beyond the purposes of this Article. For two principled approaches to criminalization, see JONATHAN SCHONSHECK, ON CRIMINALIZATION (Alan Mabe et al. eds., 1994), and DOUGLAS HUSAK, OVERCRIMINALIZATION (2008). Jonathan Schonsheck suggests three elements to limit the state's power to enact criminal legislation: the principle filter, the presumptions filter, and the pragmatics filter. *See* SCHONSHECK, *supra*, at 25. Douglas Husak suggests a structured approach to criminalization, relying on internal and external constraints. *See* HUSAK, *supra*, at 145. For a full taxonomy of criminalization theories, see ELDAR HABER, CRIMINAL COPYRIGHT 217–56 (2018).

79. Academic debate on Samaritan laws is mostly divided between those who view a duty to rescue or report as a moral duty and those who view it as an act of charity. *See, e.g.*, Anthony D'Amato, *The "Bad Samaritan" Paradigm*, 70 NW. U. L. REV. 798, 798 (1976) (stating that this area of law is unsettling because of the disparity between rule of law and morality); Weinrib, *supra* note 14, at 267 (stating although there may be a duty to give charity, the omission of any action cannot be considered a legal wrong); Yeager, *supra* note 25, at 4 (discussing how scholars viewed the duty to rescue as a duty, not charity); Heyman, *supra* note 58, at 680 (stating that a duty to rescue is "an obligation not only owed to the community itself but also to other members of that community"); Bhargava Ray, *supra* note 58, at 630–32 (discussing the conflict between freedom of choice and whether or not an individual has a moral obligation to act). For an economic analysis of altruism in the context of Samaritan duties, see Landes & Posner, *supra* note 64.

80. *See* Volokh, *supra* note 72, at 113 (explaining enforceable law enforcement policies that may lead to more reporting, including declining to prosecute sex workers or undocumented persons who report crimes against them).

But while this debate has been ongoing for several decades, much has changed in recent years in the ways that bystanders might witness perilous situations, and more closely, criminal activities.[81] The developments of digital technology and the markets that drive them have led to a potential new form of digital bystanders.[82] To understand these new forms of bystanders and their potential challenges within the context of bad Samaritans, the following Part will be further divided into two sperate discussions: Part III.A on communications of perilous situations and those that witness them online and Part III.B on platforms that could detect perilous situations, and more specifically criminal activity, using AI technology.

III. Online and Artificial Bystanders

Digital technology has created actual and potential new forms of bystanders, which include both human and machines.[83] Within the human perspective, digital technology now enables subjecting individuals to becoming online bystanders—users who witness perilous situations in real time, like disasters or crimes being committed, or digitally contacted by those in peril or danger.[84] With some notable differences, further explored below in Part III.A, this rather new form of online bystanders could be somewhat analogized to physical bystanders of perilous situations—those who initially triggered bad Samaritan legislation.

Similarly, other technological developments might also form a new category of bystanders—platforms or devices that are able to detect, to some level of certainty, that an individual is currently at risk or that a felony is being committed.[85] These "artificial bystanders" are essentially AI-based devices and services that could surround individuals almost anywhere they

81. *See infra* notes 83–87 and accompanying text.

82. *See infra* notes 83–87 and accompanying text .

83. *See* Benzmiller, *supra* note 18, at 934 (stating that the internet and other forms of electronic communications have created infinite bystanders).

84. *See id.* at 935 ("The use of electronic communication decreases the sender's awareness of how the receivers will perceive and react to the communications.").

85. *See infra* notes 230–236 and accompanying text.

go while they constantly receive and transmit data to a third party.[86]

This Part introduces the present and (near) future of bad Samaritan duties while examining their potential social and legal implications. While the current regulatory regime was not designed to apply on remote digital bystanders, it is crucial to understand how current and future technologies, along with social and market forces, might reshape the ways in which society and policymakers react to these new forms of bystanderism.[87] As the main form of bad Samaritan law applicable to digital bystanderism is that of the duty to report,[88] simply because duty to rescue generally requires physical presence, this Part will focus on the duty to report within the context of two separate (but sometimes linked) discussions, representing current and future dilemmas: online users (online bystanders) and AI-based platforms (artificial bystanders).

A. Online Bystanderism

The internet clearly had a vast impact on how humans gain knowledge and communicate with each other. By now, many individuals in modern society are often using online services as an integral part of their lives—perhaps even as the main form of communication with others.[89] As such, online services might expose individuals to perilous situations, often violent or otherwise harmful in nature, that have occurred or are

86. *See infra* notes 230–236 and accompanying text.

87. *See infra* notes 109–112 and accompanying text.

88. As will be further showed, a duty to rescue is highly limited within the digital realm, only applicable to what will be referred to as the "primary live streamer"—those individuals that, apart from livestreaming the crime, could also potentially rescue the victim. *See* Benzmiller, *supra* note 18, at 954 ("[T]here is little a bystander can do to 'rescue' the victim of a harassing communication.").

89. Notably, in what often is termed the digital divide, there is gap between individuals in terms of access and use of communication technology, thus many individuals in society might not even use social media. For more on the digital divide, see Chris Ashworth, *Bridging the Digital Divide: How to Stop Technology Leaving Young People Behind*, GUARDIAN (July 24, 2017, 2:19 PM), https://perma.cc/MA7C-9K9X.

currently occurring somewhere around the world.[90] These internet users might, for instance, be contacted directly by someone for aid or contacted indirectly, as spectators witnessing individuals in distress or in a perilous situation via, e.g., a livestream of an ongoing crime, a car accident, or other common disasters.[91] In other words, the internet, perhaps mostly led by social media platforms, has opened a gateway of transforming individuals into a new form of bystanders—those without their physical presence where perilous situations take place.[92]

Perhaps not surprisingly, we have come to witness in recent years individuals who used social media platforms to livestream criminal activities such as rape,[93] assault,[94] abuse,[95]

90. *See infra* notes 93–98 and accompanying text.

91. *See infra* notes 93–98 and accompanying text.

92. *See* Benzmiller, *supra* note 18, at 934 ("The Internet and other forms of electronic communication also foster harmful social behavior because such technology alters the nature of social interactions.").

93. *See* Mike McPhate, *Teenager Is Accused of Livestreaming a Friend's Rape on Periscope*, N.Y. TIMES (Apr. 18, 2016), https://perma.cc/2QJD-DL2L (stating that Marina Lonina was accused of livestreaming her 17-year-old friend being raped on the Periscope app); Yaron Steinbuch, *Suspects in Livestreamed Gang Rape Are Afghan Immigrants*, N.Y. POST (Jan. 26, 2017, 12:51 PM), https://perma.cc/HRX8-HQVG (describing how a three-hour rape of a woman was livestreamed on Facebook).

94. *See* J. Weston Phippen, *The Desire to Livestream Violence*, ATLANTIC (Jan. 6, 2017), https://perma.cc/BD6X-3623 (describing the livestream of the kidnapping and torturing of a man, in which the attackers were using a knife to slice off a piece of the victim's hair, down to his scalp, all while other online participants were watching and commenting).

95. *See* Alan Yuhas, *Ohio Mother who Taped Son to Wall on Facebook Live Faces Charges*, GUARDIAN (Jan. 20, 2019, 9:04 AM), https://perma.cc/6P49-AFDK (describing how a woman from Ohio was using Facebook Live to broadcast how she taped her two-year-old son to a wall).

hate crimes,[96] murder,[97] and even mass massacres,[98] to name but a handful of examples. These so-called "livestreaming crimes" largely formed the abovementioned "online bystanders": individuals who, much like witnessing a crime or other perilous situations in the physical world, now do so digitally from afar.[99] These users might either remain passive, in the sense that they only watch the livestream, or they might become more active, e.g., comment on the stream, share it with others, or contact someone for the aid of those in peril.[100]

Naturally, these livestreaming crimes raise a host of legal questions, mainly within the context of intellectual property,[101]

96. *See* Sam Levin & Amber Jamieson, *Four Suspects Charged with Hate Crimes over Beating in Facebook Live Video*, GUARDIAN (Jan. 5, 2017, 2:22 PM), https://perma.cc/9AVY-G65S (explaining how four individuals were charged with hate crimes after they livestreamed gagging and brutally attacking a person with disability); Adi Robertson, *An Anti-Semitic Shooting in Germany Was Livestreamed on Twitch*, VERGE (Oct. 9, 2019, 1:34 PM), https://perma.cc/RJP3-TL6M (describing a terror attack outside a synagogue in Germany was livestreamed on Twitch); Dave Maclean, *Two Killed in Shooting During a Livestreamed Church Service in Texas*, INDEPENDENT (Dec. 29, 2019, 8:14 PM), https://perma.cc/D992-6Z93 (describing a man who opened fire during a livestreamed church service in Fort Worth, Texas, while killing another individual aside from himself).

97. *See* Patpicha Tanakasempipat & Panarat Thepgumpanat, *Thai Man Broadcasts Baby Daughter's Murder Live on Facebook*, REUTERS (Apr. 25, 2017, 7:47 AM), https://perma.cc/928B-G2SD (reporting that a man filmed himself on Facebook killing his eleven-month-old daughter, and eventually committed suicide).

98. *See* Kevin Roose, *A Mass Murder of, and for, the Internet*, N.Y. TIMES (Mar. 15, 2019), https://perma.cc/H9D4-ML9F (describing a live video stream of mass shooting in Christchurch, New Zealand, where fifty-one people were killed and forty-eight were injured during an attack on two mosques).

99. *See supra* notes 93–98 and accompanying text.

100. *See supra* notes 93–98 and accompanying text.

101. Copyright law generally grants protection for works that are both original and fixed in a tangible medium of expression, regardless of content. *See* 17 U.S.C. § 102(a) ("Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device . . . ."). In other words, these livestreaming crimes might enjoy copyright protection, but such protection might be undesirable for society for various reasons. For more on the copyright aspects of livestreaming crimes, see generally Eldar Haber, *Copyrighted Crimes: The Copyrightability of Illegal Works*, 16 YALE J.L. & TECH 454 (2014).

and privacy rights.[102] In the context of Samaritan laws, livestreaming crimes raises crucial questions as to the legal duty of individuals in society to assist those in distress or in perilous situations by reporting it, much like the legal obligations that many states have imposed on individuals physically present at these events. In other words, the normative question that unfolds here is whether the current Samaritan duties could or should be imposed on online bystanders, i.e., does the law currently hold these users accountable for being bad Samaritans, and even if not, should reporting perilous situations online, and more specifically, livestreaming crimes, remain a moral choice or become a legal duty?

To answer this complex question, the next Part differentiates between two forms of online bystanders: communicators and viewers.[103] Online communicators are those who are either the primary livestreaming users—those communicating the perilous situation to other individuals online and physically present at the scene—or secondary users, likely not physically present at the scene but who further communicate (or share) the event to other users, hence potentially making it more visible.[104] In contrast, online viewers

102. One issue, beyond the scope of this Article, is that of the right to privacy in public places. For more on privacy in public places in the context of livestreaming, see generally Daxton R. "Chip" Stewart & Jeremy Littau, *Up, Periscope: Mobile Streaming Video Technologies, Privacy in Public, and the Right to Record*, 93 JOURNALISM & MASS COMM. Q. 312 (2016). For other legal concerns, see Kerry O'Shea Gorgone, *Live Streaming Video: Is It Legal?*, HUFFPOST (Aug. 30, 2017, 11:36 AM), https://perma.cc/5FAF-64CM; Sara Hawkins, *Legal Ins and Outs of Live Streaming in Public*, SARA F. HAWKINS (Oct. 5, 2015), https://perma.cc/8YTD-Y3Y6.

103. *See infra* Part III.B.

104. *See supra* notes 93–98 and accompanying text. Notably, and as will be further discussed, this act might also raise awareness, and some social media platform will often remove this content when they become aware of it. Facebook, for instance, claims that it removes content that expresses support or praise for organizations or individuals involved in "terrorist activity, organized hate, mass murder (including attempts) or multiple murder, human trafficking, and organized violence or criminal activity." *See Community Standards*, FACEBOOK (2020), https://perma.cc/L9ZD-AHUG.

are passive.[105] They watch the communication without further communicating or distributing it to others.[106]

1. Online Communicators

Depending on the jurisdiction, it is rather evident that the primary livestreaming user of perilous situations might face liability in the context of bad Samaritan laws. The reason is self-evident: if these individuals either have a duty to report (or to rescue) under the state Samaritan law where the event took place, and they are physically present there, they are obliged to meet the legal requirements much like any other bystander.[107] The more relevant question in this context is whether communicating the perilous situation could be considered reporting the event, thus satisfying the requirement to actively report the perilous situation.[108]

To contextualize this argument, consider the current legal requirements of selected bad Samaritan laws. Minnesota's and Hawaii's bad Samaritan laws require giving victims "reasonable assistance," which "may include obtaining or attempting to obtain aid from law enforcement or medical personnel."[109] Vermont requires one to give "reasonable assistance or care" but only if such assistance or care is not being provided by others.[110]

---

105. *See* Woodrow Hartzog, *The New Price to Play: Are Passive Online Media Users Bound by Terms of Use*, 15 COMM. L. & POL'Y 405, 411 (2010) (defining "passive online media user" as any reader, viewer, listener who makes use of a website for informational, research or entertainment purposes only, without contributing content or otherwise interacting).

106. *See id.*

107. *See infra* notes 109–112 and accompanying text.

108. *See supra* notes 33–35 and accompanying text.

109. *See* MINN. STAT. § 604A.01 (2020) ("A person at the scene of an emergency who knows that another person is exposed to or has suffered grave physical harm shall, to the extent that the person can do so without danger or peril to self or others, give reasonable assistance to the exposed person."); HAW. REV. STAT. § 663-1.6 (2019) ("Any person at the scene of a crime who knows that a victim of the crime is suffering from serious physical harm shall obtain or attempt to obtain aid from law enforcement or medical personnel if the person can do so without danger or peril to any person.").

110. *See* VT. STAT. ANN. tit. 12, § 519 (2019)

Wisconsin's law applies solely to criminal conduct and obliges one to "summon law enforcement officers or other assistance or shall provide assistance to the victim."[111] As a final example, Rhode Island simply states that one must give reasonable assistance to victims, without providing further guidance on the interpretation of the statute.[112]

Does the act of communication, by itself, currently fall under the legal requirements of assistance or reporting, thus granting live-streamers immunity for not physically aiding those in peril? Naturally, it is difficult to assess whether streaming the perilous situation—thereby communicating the event to others seeking their aid in contacting law enforcement agencies—would fall into the various vague requirements set by these different statutes. One might consider that untangling the conundrum of whether to interpret the action of livestreaming as assistance will depend greatly on the emphasis one places on the cause or the effect of such communication.[113] The *cause* behind the livestream relates to the mental element behind the communicator's intention, his *mens rea*.[114] On the other hand, the *effect* of the communication disregards mental state, focusing only on whether the livestream directly led to assistance.[115]

> A person who knows that another is exposed to grave physical harm shall, to the extent that the same can be rendered without danger or peril to himself or herself or without interference with important duties owed to others, give reasonable assistance to the exposed person unless that assistance or care is being provided by others.

111. *See* WIS. STAT. § 940.34(2)(a) (2020) ("Any person who knows that a crime is being committed and that a victim is exposed to bodily harm shall summon law enforcement officers or other assistance or shall provide assistance to the victim.").

112. *See* 11 R.I. GEN. LAWS ANN. § 11-56-1 (2020)

> Any person at the scene of an emergency who knows that another person is exposed to, or has suffered, grave physical harm shall, to the extent that he or she can do so without danger or peril to himself or herself or to others, give reasonable assistance to the exposed person.

113. *See supra* notes 93–98 and accompanying text.

114. *See supra* notes 93–98 and accompanying text.

115. *See supra* notes 93–98 and accompanying text.

Those who acknowledge the dominant rationale behind bad Samaritan laws, that is, to morally educate members of society, will vote for placing the emphasis on the cause of livestreaming crimes.[116] Those who place importance on the effects side of this debate, might be more pragmatic and reach an opposite conclusion—that the mental state is irrelevant, as long as those in perilous situations benefit from assistance prompted by the livestream.[117]

This debate is linked also to the second type of online communicators—secondary users—those not physically present at the perilous situation but who communicate the event to other users online, hence making it more visible.[118] Unlike the primary livestreamer, secondary users are in no physical position to rescue the victim but are potentially able to aid them by reporting the incident.[119] As Part III.A.2 further discusses the passive aspect of online communicators, this Part focuses on their active conduct, i.e., whether users who have shared, commented, liked, or made the livestream otherwise more visible to others, could and should bear legal liability for their actions. But as discussion on the active role that secondary users play is generally akin to the general benefits and drawbacks of the primary livestreamer (aside from not being able to grant physical assistance) and the interpretation of current bad Samaritan statutes that require some form of reasonable assistance by bystanders, they will be jointly analyzed.

Currently, courts will not likely interpret livestreaming as assistance under present bad Samaritan legislation, as it was born out of the cause rational, i.e., the mental state of the

116. *See supra* notes 61–62 and accompanying text.

117. *See* Jillian C. York, *The Murky Ethics of Facebook Live and Filming People Without Their Consent*, QUARTZ (Apr. 3, 2017), https://perma.cc/PC4E-89GA (noting that in the United States photojournalists "often live by the mantra of asking forgiveness, rather than permission" to film perilous situations).

118. *See supra* notes 93–98 and accompanying text.

119. *See* Perrin-Smith Vance, *supra* note 19, at 148 (explaining that a duty to report crimes would allow law enforcement to respond to crimes quickly and prevent leads from becoming stale). *But see* Benzmiller, *supra* note 18, at 947 ("A duty to report a crime may unacceptably violate the victim's privacy, especially for victims of sexual offenses.").

Samaritan which sparked a public outcry.[120] Still, a growth in the practice of livestreaming crimes or other perilous situations[121]—along with evidence that these livestreams might be proven highly beneficial for those in peril—can change the reasoning behind new bad Samaritan laws, thereby extending legal immunity to online communications. In other words, even if current interpretation of bad Samaritan laws will lead to holding online communicators legally liable, the normative question is whether the law *should* include livestreams as "assistance" or reporting.

To answer this question, one must first evaluate whether livestreaming of perilous situations, and perhaps mostly livestreaming crimes, is a socially desirable conduct that should be promoted by the law or by any other means. On its merits, one might argue that livestreaming crimes or other perilous situations is a socially preferable conduct as these livestreams could, *inter alia*, aid those in peril by drawing attention to the event in real time, thus increasing the plausibility that someone, including law enforcement agents, will come to their aid.[122] Communicating perilous situations might also aid individuals close to the event to be wary of it, thus raising awareness of risks in general, allowing them to avoid the area,[123] or even moving them to assist those in peril.[124]

---

120. For a discussion on present bad Samaritan legislation, see *supra* notes 109–112 and accompanying text.

121. *See* David Glance, *As Live Streaming Murder Becomes the New Normal Online, Can Social Media Be Saved?*, PHYS.ORG, (Apr. 19, 2017), https://perma.cc/UL8M-F63S (explaining that livestreaming of murders and suicides have become increasingly frequent occurrences).

122. *See* Tracy Moore, *The Benefits—and Drawbacks—of Livestreaming Crime*, VOCATIVE (Apr. 21, 2016, 1:40 PM), https://perma.cc/5H8P-GJUJ (stating that a livestreamer's attorney argued that she was streaming the assault in order to gather evidence).

123. Vigilante, later rebranded as "Citizen", is an example of a "[l]ive crime alert app" which "sends users a push notification of crimes in progress so they can either choose to avoid the area or go to the scene to broadcast live video as the crime takes place." Khari Johnson, *Apple Removes Vigilante Live Crime Alert Tool from the App Store*, VENTUREBEAT (Nov. 1, 2016, 8:18 AM), https://perma.cc/3NH8-NVRV.

124. *See* Olivia Solon, *Crime-Reporting App Vigilante Kicked Off App Store over Apple's Content Concerns*, GUARDIAN (Nov. 1, 2016, 7:05 PM),

The livestreaming of crimes specifically could aid in the fight against crimes and injustice, and perhaps even aid in strengthening the connection and trust within the community.[125] Potentially, in this context, livestreaming crimes could even strengthen the community-police relationship, as long as the latter will have sufficient digital presence and responsiveness so that the former could easily alert them.[126] It thus potentially creates a new public safety resource, prompted by the online environment.[127]

Streaming and capturing perilous situations also possess social and legal benefits beyond the immediate aid to those in peril, especially when it comes to criminal conduct.[128] They could potentially aid law enforcement agencies and victims in obtaining evidence and are thus important. This is especially true when leads to the suspect of the crime are considered "hot,"[129] or the leads provide assistance to victims (and the state) in future related legal proceedings.[130] Streaming and capturing thus hold significance for society as a whole and the victim

---

https://perma.cc/ZQ9U-6PML (indicating that that the app was designed to empower people and restore faith in law enforcement).

125. *See id.* ("[T]ransparency is the single most powerful tool against crime and injustice."). Livestreaming could, for instance, aid in documenting and reducing police brutality. *See id.* ("On one hand Vigilante talks about restoring trust between law enforcement and the community, which suggests that video streaming could help document and prevent police brutality.").

126. *See* SOCIAL MEDIA STRATEGY IN POLICING 93–94 (Babak Akghar et al., eds., 2019) ("What becomes noticeable through researching these nuanced methods is law enforcement's solicitation of remoulding the bystander effect by innovative technological means encouraging the transition of the public from passive spectators of crime to active witnesses, recognising their potential in aiding investigations.").

127. *See id.* at 94 ("These examples express an intriguing development in understanding public crime reporting as a phenomenon that provides opportunities for law enforcement.").

128. *See id.* ("They further display growing law enforcement efforts to capitalise on changing crime reporting behaviours to benefit intelligence gathering and investigations.").

129. *See* Perrin-Smith Vance, *supra* note 19, at 148 ("By compelling witnesses to report crimes as soon as reasonably possible, authorities will be able to investigate leads while they are still 'hot.'").

130. *See* SOCIAL MEDIA STRATEGY IN POLICING, *supra* note 126, at 93 (discussing how data provided by public crime reporting on social media can be used by law enforcement agencies to aid in investigations).

specifically, in both criminal and civil cases.[131] One might argue that as collecting evidence should be generally encouraged,[132] they should be promoted at least in those situations where the communicator was clearly unable to reasonably and without self-risk or harm stop the crime and notify law enforcement agencies prior or during the livestream.[133] This benefit broadens the potential benefits of kinetic bystanderism beyond the classical rationales of bad Samaritan laws, as the potential evidence gathered will most likely prove more efficient than eye-witnessed crimes in the physical realm.[134]

The negative aspects of the livestreaming of perilous situations, especially livestreaming crimes, are non-negligible.[135] For one, it would be highly difficult to evaluate what better serves society: attempting to aid those in peril personally while directly notifying law enforcement agencies or communicating the event for purposes of both potential aid and evidence. One must also bear in mind that there are several other negative consequences of such communications that must

131. These videos could eventually provide evidence to be used when prosecuting a culprit, but also, to clear suspects. *See* Ryan Tarinelli, *Bystander Uses Facebook Live to Show Little Rock Shooting Scene; Live Streams New Trend*, ARK. DEMOCRAT-GAZETTE (Mar. 20, 2017, 4:30 AM), https://perma.cc/Z9YL-FJMB ("Recording the scene can provide evidence that can be used to prosecute a culprit, he said, and can also be used to clear a defendant of any wrongdoing.").

132. *See* Moore, *supra* note 122 ("If you were witnessing a crime and unable to stop it by yourself, taking or collecting evidence or livestreaming with hopes that someone would see it and stop it would be something we'd want to encourage, and as a general matter, would be protected.").

133. *See supra* notes 126–128 and accompanying text; *see also* Moore, *supra* note 122 ("Now, if facts were to show it were done as part of a broader criminal act—if she were participating in a conspiracy or engaged in nonconsensual videotaping of someone who was currently the victim of a crime, then that act could be criminalized. Intent matters a lot." (internal quotation omitted)).

134. *See* Perrin-Smith Vance, *supra* note 19, at 148 ("[Use of livestreaming evidence] would conceivably lead to more arrests and convictions of violent felons.").

135. *See, e.g.*, Moore, *supra* note 122 (indicating that livestreaming crime raises privacy concerns for victims of crime and may lead bystanders to record an incident rather than render assistance).

be taken into account.[136] These include the privacy rights of the those in peril, even if at public places;[137] the emotional and psychological effects of such online exposure;[138] the effect of the violent or otherwise harmful nature of many livestreams on some viewers;[139] the inclusion of innocent people nearby;[140] the possibility that the livestream might be false or blown out of proportion and cause unnecessary panic due to the availability heuristic;[141] the risk that attention of users might provide tacit support for an immoral conduct;[142] the risk that it could perpetuate biases and racial profiling in the sense that the community will use discretion on which events they choose to

136. *See* Lawrence W. Sherman, *The Influence of Criminology on Criminal Law: Evaluating Arrests for Misdemeanor Domestic Violence*, 83 J. CRIM. L. & CRIMINOLOGY 1, 14 (1993) (explaining that police intervention does not always diffuse violent situations and are often less successful in domestic violence cases, with the violence escalating in retaliation for the arrest).

137. The Fourth Amendment's "reasonable expectation of privacy" standard does not generally cover public places. *Katz v. United States*, 389 U.S. 347, 351 (1967) ("What a person knowingly exposes to the public . . . is not a subject of Fourth Amendment protection."). In some instances, however, courts have held that in places designed to ensure privacy, like a telephone booth or public restrooms, the Fourth Amendment protection might apply. *See id.* at 352; U.S. CONST. amend. IV. For more on these cases and on privacy in public places, see generally Helen Nissenbaum, *Toward an Approach to Privacy in Public: Challenges of Information Technology*, 7 ETHICS & BEHAV. 207 (1997); Joel R. Reidenberg, *Privacy in Public*, 69 U. MIAMI L. REV. 141 (2014); *supra* note 102 and accompanying text.

138. *See generally* Anthony Feinstein et al., *Witnessing Images of Extreme Violence: A Psychological Study of Journalists in the Newsroom*, 6 J. ROYAL SOC'Y MED. 1 (2014).

139. *See generally id.*

140. *See* Jillian C. York, *The Murky Ethics of Facebook Live and Filming People Without Their Consent*, QUARTZ (Apr. 3, 2017), https://perma.cc/PC4E-89GA (discussing the potential risk of retaliation when individuals record other unassuming individuals without asking permission).

141. The availability heuristic suggests that individuals make mental shortcuts and heavily weigh their judgments (i.e., the plausibility of these events to reoccur) toward more recent information. *See* ANTHONY ESGATE ET AL., AN INTRODUCTION TO APPLIED COGNITIVE PSYCHOLOGY 201–02 (2005).

142. *See* Benzmiller, *supra* note 18, at 932 ("Bystanders, by their mere presence and attention, provide at least tacit support so that the bully does not act entirely alone.").

livestream or report;[143] among other potential drawbacks.[144] In the specific context of secondary users, one might argue that drawing attention to the livestream might inflict further harm on the victim—which is morally distinct from simply not aiding.[145]

Notably, this discussion extends far beyond bad Samaritan laws, as the fundamental question behind it relates to other normative and social values.[146] The answer to this conundrum will depend on whether the act of communication could lead to aiding victims or others in perilous situations, balanced against the negative impacts of these livestreams on some individuals, and on society as a whole. Imposing criminal liability on either form of communicators might potentially lead to a reduction in livestreams of perilous situations like livestreaming crimes—lowering the incentive for at least some online communicators to livestream and abstain from aiding those in peril otherwise.[147] On the other hand, while there are many drawbacks to these livestreams, they might still aid those in

143. *See* Scott Lucas, *The Citizen's Dilemma*, S.F. MAG. (Dec. 20. 2017), https://perma.cc/7K7H-Y9GZ (explaining that when individuals hastily report on their community members it encourages bias); Taylor Hatmaker, *Citizen Expands Its Crime-Tracking Alert App to Baltimore*, TECHCRUNCH (Feb. 15, 2019, 7:51 PM), https://perma.cc/E849-7RDW ("[T]here's no evidence this state of hyper-awareness does any quantifiable good, and at least some evidence that it can actually put people, specifically people of color, at *more* risk due to implicit bias and racial profiling.").

144. For more on racial biases in the context of bad Samaritan laws, see Swan, *supra* note 34, at 1044–45.

145. *See* Bhargava Ray, *supra* note 58, at 630 (arguing that the law should not punish the mere failure to act, even if such a rule seems heartless or immoral because an individual does not violate another individual's rights simply by declining to aid them).

146. *See generally* Joel Feinberg, *The Moral and Legal Responsibility of the Bad Samaritan*, 3 CRIM. JUST. ETHICS 56 (1984) (discussing moral values and objections to bad Samaritan statutes).

147. *See* Benzmiller, *supra* note 18, at 961 ("Many people refuse to assist victims of crime out of fear of retribution; if reporting is to be a viable alternative for these individuals, they must believe that police will respect their anonymity.").

peril and in obtaining evidence,[148] and thus any legal obligation in this context must be carefully examined.

Still, the law might be highly limited in regulating these livestreams if policymakers focus on online communicators—and more specifically, the primary livestreamers. While communicators are the initial cause of this potential problem, bad Samaritan laws will not likely serve as a deterrent to livestream these events due, *inter alia*, to the fact that many of these livestreaming crimes were communicated by those committing the crime.[149] The more fundamental and applicable question in the context of Samaritan duties is that of passive online users who view the livestream but do not actively react to it or actively participate in the dissemination of the content in any way.

2. Online Viewers

Do current bad Samaritan laws impose a legal duty of assistance on passive online viewers? To understand the potential legal liability that online viewers might currently face, it is crucial to interpret the general duty-to-report requirements. But as duties to report crimes and other perilous situations vary between states, each potentially leading to somewhat of a different interpretation, this Part explores potential legal lability through scrutinizing the example of bad Samaritans laws in the state of Texas ("Texas example").

The Texas Penal Code deems it an offense if the observer of a felony "in which a reasonable person would believe that an offense had been committed in which serious bodily injury or

148. *See* SOCIAL MEDIA STRATEGY IN POLICING, *supra* note 126, at 94 (explaining that when social media is a platform where crime reports are shared, law enforcement can capitalize on changing crime reporting behaviors to benefit intelligence gathering and investigations).

149. *See* Olivia Solon, *Why a Rising Number of Criminals Are Using Facebook Live to Film Their Acts*, GUARDIAN (Jan. 27, 2017, 5:33 PM), https://perma.cc/N7Y4-QY9T (last updated Sept. 20, 2017, 12:08 PM) (stating that the practice of documenting one's crimes is on the rise even though it is self-incriminating). This argument further raises an odd, but perhaps rather self-evident, question: can those who committed a crime be liable also for not aiding their own victim, thereby potentially breaking the bad Samaritan law in their jurisdiction?

death may have resulted" did not immediately report it to a law enforcement agency, and as long as a reasonable person "would believe that the commission of the offense had not been reported" and as long as the reporting does not place them in "danger of suffering serious bodily injury or death."[150] While this example will mainly serve to discuss livestreaming crimes, as it is confined to felonies, the discussion of non-criminal perilous situations that could be applicable under other statutes would similarly apply in most instances.

The Texas example does not currently differentiate between offline and online observers of a felony; thus, it could theoretically be applied to online viewers.[151] But unlike the discussion of online communicators, digital viewership might change the interpretation of current (and desirable) bad Samaritan duties, regardless of examining whether these viewers rendered assistance or not. In other words, the initial issue here would be the potential changes in the mental state of the viewer which might be different in digital viewing. Factually, as mental state is already embedded within the criminal system,[152] the outcome of such analysis will most likely depend on specific events, along with many potential barriers that will be further explored.

The broader question is thus not whether the law could impose affirmative duties on online users to assist those in perilous situations, but whether it should. The answer relates, *inter alia*, to the potential differences between offline and online observers, i.e., that one must scrutinize if observing a felony changes due to the medium through which the felony is communicated.[153] In other words, the fundamental normative question is whether witnessing a felony live online (or more generally, perilous situations in some states) is inherently different than doing so in the physical world, and whether the potential differences between the two conducts should play a

150. TEX. PENAL CODE ANN. § 38.171 (West 2020).

151. *See id.* (explaining the Texas bystander statute).

152. *See* Francis Bowes Sayre, *Mens Rea*, 45 HARV. L. REV. 974, 974 (1932) ("No problem of criminal law is of more fundamental importance or has proved more baffling through the centuries than the determination of the precise mental element or *mens rea* necessary for crime.").

153. *See supra* notes 105–108 and accompanying text.

role in either interpreting the Samaritan duty to report or creating a new one.

The normative aspects of imposing bad Samaritan duties to report do not generally change in light of technology.[154] On the contrary, duties to report perilous situations like crimes should even be enhanced when online, because unlike in the physical world, online bystanders are exposed to no immediate risk for aiding those in peril.[155] Factually, in many instances users who reported these events were directly responsible for leading law enforcement agencies to the responsible parties.[156] Under this argument, policymakers must impose even stricter duties to report on online bystanders than those imposed in the physical world, as these users could potentially aid those in distress or perilous situations without the risks associated with reporting them when physically present. Under this rationale, there is little support for not imposing similar duties simply because the medium of viewership changes.

The barriers for imposing such affirmative duties are both pragmatic and normative in nature. On the pragmatic side, one challenge is that of the mental state and cognitive failures of online bystanders.[157] Unlike witnessing an event in the physical world, viewing a livestream of a perilous situation is highly subjective to the viewer, in the sense that the communicated scene might be perceived as either a prank or something that is seemingly blown out of proportion, to name but a few potential cognitive interpretations of online viewership.[158] These viewers might not be able to distinguish between an actual crime being

154. *See supra* notes 55–61 and accompanying text.

155. *See* TEX. PENAL CODE ANN. § 38.171 (indicating that in the Texas example the duty to report only attaches as long as the reporting does not place the individual in "danger of suffering serious bodily injury or death").

156. *See, e.g.*, Kelly Malone, *Live Streaming Crime: How Do We Police the Internet?*, CBC NEWS (Apr. 5, 2017, 1:10 PM), https://perma.cc/M57C-HTEQ (explaining that officers arrested three people after someone watching the weekend live stream reported the assault to police).

157. *See supra* notes 151–152 and accompanying text.

158. *See* Stuart Wolpert, *UCLA Psychology Study Explains When and Why Bystanders Intervene in Cyberbullying*, UCLA NEWSROOM (Jan. 14, 2016), https://perma.cc/YY7Q-SBKX (discussing an online study conducted in which participants' responses to cyberbullying varied based on the level of personal expression of the victim's feeling).

committed or something else. They might not be able to tell if the event is occurring live and how much the source, which they might not be familiar with at all, is in fact reliable. Consider again in this context the Texas example, requiring that "a reasonable person would believe that an offense had been committed."[159] How should the law in Texas interpret such a subjective requirement given the aforementioned cognitive biases and challenges?

Without diminishing the importance of such biases and pragmatic aspects, suppose that this challenge is somehow met, meaning that viewers comprehend that a crime or another perilous situation is being conducted and communicated in real time. If bad Samaritan laws were initially drafted as a response to the public's outcry against humans not aiding those in need,[160] then should it be different where dozens of people witness a livestream of, say, a fifteen-year-old girl getting raped and do nothing?[161] What about livestreaming an ongoing fatal car accident or common disaster with hundreds of individuals in

159. TEX. PENAL CODE ANN. § 38.171 (West 2020).

160. As mentioned, bad Samaritan laws were often triggered by a horrific event, by which bystanders chose not to aid or report law enforcement agencies. One famous incident occurred in Las Vegas, Nevada, where David Cash witnessed the attack of a seven-year-old girl in a restroom by his friend Jeremy Strohmeyer and did not aid the girl or report the event. *See* Jeremy Waldron, *On the Road: Good Samaritans and Compelling Duties*, 40 SANTA CLARA L. REV. 1053, 1054 (2000) (recounting the event). The victim, Sherrice Iverson, was then sexually molested and strangled to death. *Id.* After knowledge of this event, many have proposed to legislate both state and federal bad Samaritan laws. *See id.* at 1055 (attributing pressure for a Nevada Good Samaritan law to the event); Perrin-Smith Vance, *supra* note 19, at 135, 140 (noting that the Nevada incident led to the enactment of bad Samaritan laws). The death of Princess Diana sparked public discussion as well, because photographers who witnessed the crash did not offer assistance to the victims. *See id.* at 140. Vermont passed its "Good Samaritan Statute" in response to the death of a young woman in New York, which occurred while her nearby neighbors who witnessed the crime did not report it. *Id.*

161. This event refers to an incident which occurred in March 2017 in Chicago. There, a Facebook live stream showed the gang rape of a fifteen-year-old girl, and the forty people who watched the livestream did not report the crime. *See* Malone, *supra* note 156. One response is that regardless of the venue—physical or digital—it is generally undesirable to enact laws due to specific events. *See* Givelber, *supra* note 13, at 3171.

need of aid, while users carelessly watch the livestream and do nothing?

Notably, there are a few other important differences between physical and online viewership. One main difference is that of an "online disinhibition effect,"[162] whereby users' inhibitions, often present in the physical world, are somewhat loosened online.[163] Online viewership could lead to desensitization of users, thereby making them less likely to feel obligated to report these events.[164] One main example of such effect is users' experience of "dissociative imagination"—essentially the feeling of escapism that the online world could create (i.e., that people often perceive online interaction as "less real")—combined also with dissociative anonymity and perceived invisibility, leads them to act differently than in the physical world.[165]

---

162. The online disinhibition effect relates to the "loosening of social restrictions and inhibitions that are normally present in face-to-face interactions." *Online Disinhibition Effect (Suler)*, LEARNING THEORIES (Dec. 15, 2015), https://perma.cc/6C2S-UB6E; *see generally* Laura Martocci, *Livestreamed Violent Criminal Acts*, PSYCH. TODAY (June 9, 2017), https://perma.cc/Q68K-N3CG (discussing how online disinhibition applies in cases of terrible crimes being streamed). John Suler further divides the effect into two categories: benign and toxic disinhibition. *See generally* John Suler, *The Online Disinhibition Effect*, 7 CYBERPSYCHOLOGY & BEHAV. 321 (2004) (explaining how benign disinhibition might lead a person to be more genuine over the internet than in their normal life, whereas toxic disinhibition results in negative behaviors that would be avoided in the real world).

163. *See* Martocci, *supra* note 162 (defining the online disinhibition effect).

164. *See id.* (linking the online disinhibition effect to desensitization towards acts of violence).

165. *See Online Disinhibition Effect (Suler)*, *supra* note 162, at 321 (identifying differences in behaviors online and in the physical world as online disinhibition). Online disinhibition is due to several reasons: (1) dissociative anonymity, whereby anonymity is translated into people feeling safer and protected to act online, and thereby could be more engaging in antisocial or otherwise harmful behavior (or less inhibited by social conventions and restraints); (2) invisibility that enables lowering inhibitions and allowing misrepresentation; (3) asynchronicity, whereby many online actions are not occurring in "real time," enabling users to communicate when they desire; (4) solipsistic introjection, whereby users might feel comfortable communicating with other users as they assign "imagined characteristics to another person based off of their messages and online persona;" (5) dissociative imagination, meaning that users might view online communication as some sort of a game

In general, viewers might still not feel obligated to aid others in peril, whereas the event seems detached from their lives, once again due to various cognitive biases, like that of the so-called "bystander effect," as they might assume that someone else, perhaps even the platform, will take action;[166] and that unlike with physical presence, they have limited ways to actually aid.[167] Notably, however, this bystander effect might not be fully translated into the digital era, as there is evidence to support the conclusion that people tend to behave kindly toward strangers on the internet more so than in real life situations.[168]

Going back to the Texas example, one must consider whether "a reasonable person would believe that the commission of the offense had not been reported."[169] Physical viewers might be better positioned to make such a judgement, as online viewers have little knowledge of other individuals present at the event, let alone if other users have reported it

---

that the physical rules do not apply; and (6) minimization of status and authority, in so much as the traditional forms of authority that are often translated into dress, body language, name titles, and their environments, do not occur online, letting users treat everyone as other users or peers. *See id.* at 321 (listing Suler's causes for online disinhibition); *id.* at 322–24 (articulating the reasons for online disinhibition); *see also* Darby Dickerson, *Cyberbullies on Campus*, 37 U. TOL. L. REV. 51, 62–63 (2005) (describing the drawbacks of anonymity in the context of cyberbullying).

166. Users might be reluctant to report these crimes due to a bystander effect, first discovered and coined by Latané & Darley—whereby the presence of other bystanders might reduce the likelihood of bystanders to intervene, because people often interpret the situation as less-serious, and due to diffusion of responsibility, whereby people conceive that the responsibility for action is diffused between them, thus they feel less guilty or responsible for aiding. *See* Bibb Latané & John M. Darley, *Bystander "Apathy"*, 57 AM. SCIENTIST 244, 248–49, 265–66 (1969) (describing the bystander effect); Swan, *supra* note 34, at 984–86 (examining the bystander effect and elaborating on diffusion of responsibility). Others, however, questioned the extent to which the bystander effect occurs online and argued that diffusion of responsibility could work both ways on the internet due, *inter alia*, to ambiguousness in assessing the group size. *See* WALLACE, *supra* note 34, at 198.

167. *See* Rossalyn Warren, *When Rape is Broadcast Live on the Internet*, BUZZFEED (Apr. 20, 2016), https://perma.cc/5M46-AKBL (noting that viewers of a livestreamed rape, who did not know how to intervene, were frustrated).

168. *See* WALLACE, *supra* note 34, at 192.

169. *See* TEX. PENAL CODE ANN. § 38.171 (West 2020).

prior to them.[170] Thus, here too, it would be highly problematic to interpret the duty to report requirements to apply online or impose it by new legislation.

Another challenge relates to the limits of the law to regulate such behavior online. These livestreams could often occur in remote places, thus raising various legal challenges.[171] One major challenge is jurisdictional in scope: as bad Samaritan laws vary greatly between states, and are often even entirely absent in some jurisdictions, it would be highly difficult to evaluate if the duty should be imposed at all and on whom.[172] In other words, unlike the physical world, the online bystander (i.e., the viewer) might be located in a different legal jurisdiction from where the crime takes place.[173] In this instance, supposing that the law could be interpreted to apply, which bad Samaritan law would be invoked? The legal jurisdiction where the perilous situation occurs (and communicated from)? Or the legal jurisdiction where the viewer of it is currently present? In other words, online viewership in the context of bad Samaritan laws raises questions of jurisdiction, which have often been raised in other contexts of online regulation.[174]

---

170. *See* Suler, *supra* note 162, at 322 (acknowledging the anonymity of agents over the internet).

171. *See Facebook Statistics and Facts*, MARKET.US, https://perma.cc/B3AB-SE9L (last updated Aug. 4, 2020) (demonstrating that Facebook, a platform that facilitates livestreaming, has billions of active users from all over the world).

172. *Compare* TEX. PENAL CODE ANN. § 38.171 (requiring a bystander to report a crime if a reasonable person would believe it had not yet been reported), *with* MINN. STAT. § 604A.01 (2020) (implementing a duty to assist in emergencies).

173. Arguably, this scenario could also occur in the physical realm, as an individual might be physically present in one jurisdiction, while viewing a perilous situation that occurs in another, raising jurisdictional questions of legal interpretation of the statute in question.

174. The legal debate on territorial regulations and the virtual boarders of the internet has received much scholarly attention. *See generally* David R. Johnson & David Post, *Law and Borders—the Rise of Law in Cyberspace*, 48 STAN. L. REV. 1367 (1996) (discussing uncertainty of jurisdictions given the lack of geographical borders in cyberspace); JACK GOLDSMITH & TIM WU, WHO CONTROLS THE INTERNET? ILLUSIONS OF A BORDERLESS WORLD (2006) (exploring how the internet is bordered and postulating that in some cases there is a national or jurisdictional control exercised); Jennifer Daskal, *Borders and*

Enforcement is also a substantial challenge. Aside from the aforementioned jurisdictional problem that could be challenging from an enforcement perspective, it would also be difficult for enforcement agencies to track down some users, that is, without infringing upon human rights and liberties, especially when these users could be numerous and potentially scattered all over the world.[175] Even upon detection, it will be difficult to bring them to justice while also proving their active viewing of the livestream.[176] It might also lead to biased selective enforcement against some users, and prosecutors might feel reluctant to file charges on these or other grounds.[177] Finally, given that bystanders are already rarely prosecuted in the physical world, prosecutions in the digital world will be even more challenging and unlikely.[178] This is amplified by the low deterrent value of many current penalties.[179]

These challenges are substantial from both pragmatic and legal perspectives. It is unlikely that online bystanders will play a role as legally bound Samaritans. It is also generally undesirable and seems more likely that the law will play a

Bits, 71 VAND. L. REV. 179 (2018) (exploring the jurisdictional reach of law enforcement with respect to data transmitted across borders).

175. *See* Janet Davidson, *Why It's So Hard to Catch Online Predators*, CBC NEWS (Oct. 17, 2012, 5:23 AM), https://perma.cc/B9CM-3DJS (last updated Oct. 18, 2012) (noting that online criminals can take multiple actions to preserve their anonymity, how the volume of internet activity makes identifying individual offenders difficult, and how in some cases, such as the reporting of humans rights violations, this can be a good thing).

176. *See* Malone, *supra* note 156 (raising the practical difficulties of tracking down online offenders).

177. *See* Perrin-Smith Vance, *supra* note 19, at 145–47 (highlighting the disproportionate enforcement of bad Samaritan laws when the bystander might have been involved in the crime and referencing Nevada prosecutors who feel "that individuals who fail to report will be unwilling to come forward as witnesses at a later date for fear of prosecution under the statute").

178. *See* Givelber, *supra* note 13, at 3172–95 n.213 (reporting how eight different states' bad Samaritan laws resulted in very few convictions).

179. *See* VT. STAT. ANN. tit. 12, § 519(c) (2019) (establishing that a violation of the bad Samaritan duty in Vermont cannot result in a fine greater than $100).

limited part—if any—in regulating online bad Samaritans.[180] Subjecting online viewers to affirmative duties to report will be both a constitutional and legal challenge, even if the state can assert its interest in protecting individuals from harmful activities,[181] as it will not likely be considered as narrowly tailored to achieve the state's interest, considering the existence of other potential, less intrusive measures that could better achieve this interest.[182] If people tend to aid each other in the physical world regardless of a legal obligation to do so, then perhaps some online spectators will act accordingly.[183]

Still, online platforms might play a crucial role with or without direct regulation. As these livestreams might be proven crucial for evidence, policymakers could, for instance, oblige online platforms that enable livestreaming in general to retain the data, at least for a time, even if they decide to block the livestream from the general public.[184] The problem might arise from technologies whose infrastructure purposely limits the time any streaming is accessible to other users, and hence, is

180. Notably, end-users that are not entirely passive in the livestreaming, e.g., they comment on the crime or the victim, might be subjected to various forms of liability, depending on their communication, such as defamation.

181. *See* Reno v. ACLU, 521 U.S. 844, 875 (1997) (asserting that a governmental interest in protecting children from harmful content on the internet was generally outweighed by the First Amendment rights of other internet users); Perrin-Smith Vance, *supra* note 19, at 146 (exploring how reporting requirements bring the interest of the government in protecting its citizens in direct conflict with the citizens' constitutional right to free speech through the forced speech inherent in the reporting).

182. *See* Perrin-Smith Vance, *supra* note 19, at 146–47 (illustrating that alternatives to bad Samaritan laws can be equally effective without infringing constitutional rights).

183. *See generally* Hyman, *supra* note 55 (providing evidence that while bad Samaritan laws were often enacted due to anecdotal events, people tend to attempt rescuing each other in most cases).

184. *See* Michael H. Keller, *Bipartisan Bill Targets Online Spread of Child Sex Abuse Material*, N.Y. TIMES (Dec. 10, 2019), https://perma.cc/9SX6-C8UJ (reporting on legislative efforts to lengthen the period of time that tech companies must retain data of illegal photos and videos on their platforms).

more disposable in nature.[185] But while evidence is important, it does not directly relate to bad Samaritans rationales.[186]

If policymakers insist on regulating such conduct online, then perhaps non-traditional regulatory mechanisms might prove successful. Law enforcement agencies could, for instance, strengthen their already established public-private partnerships to aid in crime detection, perhaps even creating a joint taskforce.[187] But given the potential negative impact of livestreaming crimes on individuals, the state might also decide to vote against imposing bad Samaritan duties online in general, and fight against these livestreams, while pushing towards regulating the removal of such content. In some jurisdictions, they might even turn their partnerships with the private sector to be somewhat legally mandated, thus imposing liability on social media platforms that fail to quickly remove violent material from their platforms.[188]

This debate obviously extends beyond bad Samaritans laws, as it includes the rather controversial discussion on content

185. *See* Josh Constine, *Instagram Launches Disappearing Live Video and Message*, TECHCRUNCH (Nov. 21, 2016, 10:00 AM), https://perma.cc/8BYZ-EJG8 ("Instagram Live lets you broadcast video to your followers in real-time, but they can only watch while you're still streaming. No replays.").

186. *See* Hyman, *supra* note 55, at 656 (observing that proponents of bad Samaritan laws justify them under the rationale that they would decrease bystander inaction in emergencies).

187. In Australia, for example, a social media taskforce was formed to stop the publication of violent terror content and to keep Australians safe online. *See* Media Release, Scott Morrison, *Prime Minister of Australia, Stronger Action Against Terror Content* (June 30, 2019), https://perma.cc/JQ4A-3SRQ (announcing the taskforce); *see also* Craig Timberg et al., *The New Zealand Shooting Shows How YouTube and Facebook Spread Hate and Violent Images—Yet Again*, WASH. POST (Mar. 15, 2019, 6:01 PM), https://perma.cc/BT83-9VHQ ("The department said it is working with social media platforms to remove the clips and urged the public to report objectionable content if they come across it.").

188. *See* Press Release, Christian Porter, Att'y Gen. for Austl., Tough New Laws to Protect Australians from Live-Streaming Violent Crimes (Mar. 30, 2019), https://perma.cc/C6B7-DZDW

> [T]he Criminal Code Amendment (Unlawful Showing of Abhorrent Violent Material) Bill 2019 will include new offences with penalties of up [to] 10 per cent of a company's annual turnover and potential prison sentences for executives of social media companies who fail to act to remove abhorrent violent material from their platforms.

manipulation and removal in general.[189] Currently, however, this discussion might sound redundant as online platforms in the United States are not likely to be held accountable for most of the content that is present on their platforms.[190] This is due to the prevailing interpretation of § 230 of the Communications Decency Act,[191] which exempts platforms from liability regarding third-party content that the platforms host,[192] including violent or offensive materials.[193] The general normative justification of such immunity was that in order to preserve free speech online, we must avoid viewpoint-based regulations and problems of collateral censorship, i.e., adopting intermediary liability rules that will cause over-censorship of

189. *See* Michal Lavi, *Do Platforms Kill?*, 43 HARV. J.L. & PUB. POL'Y 477, 525–43 (2020) (exploring arguments for and against imposing liability on intermediaries for undesirable content that they might host); Nellie Bowles, *The Complex Debate Over Silicon Valley's Embrace of Content Moderation*, N.Y. TIMES (June 5, 2020), https://perma.cc/X3KD-AQ5X (last updated June 30, 2020) (discussing debate over the extent to which tech companies should regulate or moderate content on their platforms).

190. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1175 (9th Cir. 2008) ("If you don't encourage illegal content, or design your website to require users to input illegal content, you will be immune.").

191. 47 U.S.C. § 230.

192. *See id.* ("No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."); Eric Goldman, *The Ten Most Important Section 230 Rulings*, 20 TUL. J. TECH. & INTELL. PROP. 1, 2–9 (2017) (recounting examples of how the courts have systematically held that this immunity stands). *But see* Safia Samee Ali, *Who Is Responsible for Stopping Livestreamed Crimes?*, NBC NEWS (Mar. 23, 2017, 2:17 PM), https://perma.cc/2DBY-FKLS (emphasizing that platforms have a legal obligation to terminate streaming and notify law enforcement of child pornography); Susan Freiwald, *Comparative Institutional Analysis in Cyberspace: The Case of Intermediary Liability for Defamation*, 14 HARV. J.L. & TECH. 569, 594–96 (2001) (reporting that initially Congressmembers wanted some requirements for platforms to screen out pornographic content); Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401, 403–23 (2017) (criticizing § 230 and suggesting that it should be adjusted through legislation or altered judicial interpretation).

193. *See* Danielle Keats Citron, *Cyber Civil Rights*, 89 B.U. L. REV. 61, 116–17 (2009) (noting sweeping immunity under § 230 even for content with a violent or offensive nature).

otherwise protected content.[194] In other words, from a legal perspective, it is currently difficult to regulate conduct like livestreaming crimes by mandating platforms to either enable or disable them.

The legal constraints do not, however, rule out the aforementioned voluntary partnership. Indeed, in some instances, platforms like Facebook and YouTube were very responsive to alerts by law enforcement agencies regarding livestreaming crimes, and quickly stopped or deleted these videos upon such alerts.[195] Some companies have even invested in automated systems to detect such undesired content,[196] hired

194. *See* Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2309–10 (2014) [hereinafter Balkin, *Old-School/New-School Speech Regulation*] (defining collateral censorship as the over-censorship of protected content incentivized by a state holding one private party liable for another party's speech); Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 U.C. DAVIS L. REV. 1149, 1176–77 (2018) [hereinafter Balkin, *Free Speech in the Algorithmic Society*] (asserting that collateral censorship is a direct corollary of intermediary liability rules and involves limiting freedom of speech)*; see generally* Alexander Tsesis, *Terrorist Speech on Social Media*, 70 VAND. L. REV. 651 (2017) (providing more information about criminal and terrorist speech in the context of social media).

195. In the case of the New Zealand shooting video on March 15, 2019, Mia Garlick, Facebook's director of policy in Australia and New Zealand, said: "New Zealand Police alerted us to a video on Facebook shortly after the livestream commenced and we quickly removed both the shooter's Facebook and Instagram accounts and the video." She further provided that: "We're also removing any praise or support for the crime and the shooter or shooters as soon as we're aware. We will continue working directly with New Zealand Police as their response and investigation continues." Jason Abbruzzese & Brandy Zadrozny, *Streamed to Facebook, Spread on YouTube: New Zealand Shooting Video Circulates Online Despite Takedowns*, NBC NEWS (Mar. 15, 2019, 11:53 AM), https://perma.cc/XUQ9-HDTU.

196. One mechanism that platforms often use in the context of copyrighted materials and pornography is creating a marked copy (a hash) and automatically blacklisting those who repost it online. This method, however, has proved ineffective in many instances. *See* Timberg et al., *supra* note 187 (explaining the hash method); Samee Ali, *supra* note 192 ("When things happen in real time, you don't know that will happen next and it's extremely difficult for automated technology to monitor live events . . . ."). In the context of AI systems that will detect livestreaming crimes, Facebook claims that it will invest $7.5 million in "new research partnerships with leading academics from three universities, designed to improve image and video analysis

extra staff to manually "monitor and promptly remove violent content,"[197] and publicly opposed the posting of such content by users.[198]

While simply examples at this point, these incidents might reveal how the concept of bad Samaritans might lose the battle for values of protecting individuals from harmful materials. Bear in mind, however, that these platforms are often reluctant to remove content as it resides within the heart of surveillance capitalism.[199] Thus, they might feel conflicted between content removal per law enforcement agencies' requests and maintaining the content to retain its profit potential.[200] If society deems the viewership of livestreaming crime as inappropriate, then social norms should attempt to shape, or even simply nudge, the market to provide appropriate tools for users to report these streams.[201] Online platforms, in turn, might react to consumers' expectations and enable, or even

technology." Jon Russell, *Facebook Introduces 'One Strike' Policy to Combat Abuse of its Livestreaming Service*, TECHCRUNCH (May 15, 2019, 2:09 AM), https://perma.cc/6SA7-Z9ZB.

197. *See* Hashela Kumarawansa, *Facebook Live Acts Against Livestreaming Crime*, SBS NEWS (Apr. 5, 2017), https://perma.cc/89JV-Q6L8.

198. *See* Samee Ali, *supra* note 192 (quoting a Facebook spokeswoman as saying "[w]e take our responsibility to keep people safe on Facebook very seriously and will remove videos that depict sexual assault and are shared to glorify violence").

199. Surveillance capitalism generally refers to the commodification of personal data for profit. *See generally* SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM (2019) (examining the concept of surveillance capitalism, how it differs from traditional capitalism, and how it should be restricted or limited). According to Becca Lewis, a researcher at Stanford and the think tank Data & Society, online companies "have a content-moderation problem that is fundamentally beyond the scale that they know how to deal with . . . . The financial incentives are in play to keep content first and monetization first." *See* Timberg et al., *supra* note 187 (quoting Lewis).

200. *See* Timberg et al., *supra* note 187 (pointing out that companies are economically incentivized not to tightly moderate content).

201. *See* Maria Shao, *Social Pressures Affect Corporate Strategy and Performance*, STAN. SCH. BUS. (Dec. 1, 2009), https://perma.cc/7U2Q-WKN8 ("Greater social pressure can result in better social performance. In other words, firms step up responsible behavior in response to pressure.").

promote, such notifications with them or with law enforcement agencies.[202]

Thus, platforms, even without direct binding regulation, could play a substantial role in regulating online Samaritans. With potential technical limitations,[203] they might, for instance, decide that such content—regardless of its potential benefits for those in peril—should not be promoted on their platforms and thus act to remove it regardless of whether it was communicated to enforcement agencies. Many websites and social platforms currently include some form of a flagging mechanism to report inappropriate or otherwise abusive content.[204] To enhance collective action and avoid the bystander effect (or syndrome)—where one might suspect that others have already flagged the event and thus be reluctant to do so—platforms could communicate to their users whether the livestream was already flagged or not.[205]

The role of platforms in the governance of behavior, further discussed in Part IV, holds great significance beyond the question of livestreaming perilous situations. But overall, the internet greatly challenges the fact that individuals might now

---

202. Ari Waldman, a leading authority on law and technology, argued that "[p]latforms that have a history of not just allowing harassment to occur, but failing to do anything when they hear about harassment, we have to make a choice to not use those platforms." *See* Malone, *supra* note 156 (quoting Ari Waldman).

203. In the livestreaming of the slaughter in two New Zealand mosques in 2019, social media platforms were highly limited in their ability to prevent the content dissemination. *See* Timberg et al., *supra* note 187 (reporting on a rapid spread of the footage that could not be stopped).

204. *See, e.g.*, *How do I Report a Live Video?*, FACEBOOK (2020), https://perma.cc/XHE4-WV4V (instructing how to report a live video on Facebook); *Report Inappropriate Content*, YOUTUBE (2020), https://perma.cc/5E6S-EJG9 (allowing for inappropriate content on YouTube to be flagged); *see also* Vanessa Callison-Burch et al., *Building a Safer Community with New Suicide Prevention Tools*, FACEBOOK (Mar. 1, 2017), https://perma.cc/YD89-2JE9 (detailing Facebook's suicide prevention tools for posts and livestreams); *How to Report Things on Facebook*, FACEBOOK (2020), https://perma.cc/357X-X8P8 (giving general information about how to report various types of content on Facebook); Kate Crawford & Tarleton Gillespie, *What Is a Flag For? Social Media Reporting Tools and the Vocabulary of Complaint*, 18 NEW MEDIA & SOC'Y 410, 411 (2016) (providing information on flagging systems in social media).

205. *See* Samee Ali, *supra* note 192.

become spectators of felonies or other perilous situations, without having a physical presence in the scene, while also potentially increasing the number of other individuals who will become such spectators due, to some extent, to their own actions.

Whether it will be policymakers, platforms, or end-users who shape and reconstruct the ways in which these livestreams will be regulated, they must also account for both the benefits and drawbacks of such technology in light of the moral and social responsibilities of society. Eventually, the answer to this conundrum will partially lie within the hands of users, deciding whether and how to report, as long as platforms enable it.[206] Notably, altruism and acts of kindness often occur online, e.g., helpful and quick replies to requests for information, in online gaming, and in emotional support forums.[207] It will also depend on online platforms, shaping how users report and where they stand on content removal in this context.[208] But the ways that users react will greatly depend on social norms regarding such livestreams and their reporting, and must include sufficient awareness that could be promoted, at least to some extent, by educational programs for users that will rely on the social duty for societal members in the online environment.[209]

Normatively, however, this question could be broader, in the sense that Samaritan laws might require reevaluating and adapting to this relatively new form of bystander that uses non-traditional measures to report crimes to the online community as a whole rather than directly to law enforcement agencies.[210] But new forms of bystanders might also begin to

206. *See* Malone, *supra* note 156 (urging people to report online crime); Pranjal Gupta et al., *Live Crime Reporting*, 5 IRJET 2927, 2928–30 (2018) (proposing an app that would streamline anonymous live-crime reporting).

207. For more on online altruism and acts of human kindness, see WALLACE, *supra* note 34, at 190–207.

208. *See supra* note 204 and accompanying text.

209. *See* Anne Kleinsasser et al., *An Online Bystander Intervention Program for the Prevention of Sexual Violence*, 4 PSYCH. VIOLENCE 227, 232–33 (2015) (detailing the success of a bystander intervention program for sexual assault prevention). *But see* Swan, *supra* note 34, at 991–94 (highlighting the limitations of bystander intervention training programs and casting skepticism towards their supposed success).

210. *See supra* notes 109–112 and accompanying text.

form—digital bystanders relying on AI technology that could aid platforms in automatically detecting potential crimes in real time. The Samaritan duties that could be imposed on these new "artificial bystanders" will be further discussed and evaluated in the following Part III.B.

B. Artificial Bystanderism

Much like almost any other law, Samaritan laws were initially crafted to regulate human conduct.[211] But technological advancement in the field of AI, where non-humans are becoming more interpretative of human conduct, including potential criminal activity, opens the floor to introduce a rather new form of Samaritans—artificial ones.[212] This new potential form of Samaritans could soon become a reality due to advancement in developments of AI and the Internet of Things (IoT), whereby regular objects (or "things") have become connected to the internet,[213] along with the increasing possibilities of these devices using sensors and advanced speech recognition or natural language processing (NLP) capabilities. Coupled with

211. *See* United States v. Wilson, 159 F.3d 280, 295 (7th Cir. 1998) (Posner, J., dissenting) ("The purpose of criminal laws is to bring about compliance with desired norms of behavior.").

212. *See* John O. McGinnis, *Accelerating AI*, 104 NW. U. L. REV. 366, 368–72 (2010) (asserting that functional AI is possible and touting the potential for AI to aid in social decision-making).

213. The term "Internet of Things" (IoT) describes devices or sensors that "connect, communicate or transmit information with or between each other through the Internet." FED. TRADE COMM'N, INTERNET OF THINGS: PRIVACY & SECURITY IN A CONNECTED WORLD 6 (2015), https://perma.cc/99HS-Z62B (PDF); *accord* Kevin Ashton, *That 'Internet of Things' Thing*, RFID J. (June 22, 2009), https://perma.cc/BE62-K8EG (encouraging the development of computers that gather information on their own). For other legal considerations that relate to IoT, see Scott R. Peppet, *Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security, and Consent*, 93 TEX. L. REV. 85, 98–117 (2014) (surveying numerous types of sensors tying objects into the IoT); see generally Eldar Haber, *Toying with Privacy: Regulating the Internet of Toys (IoToys)*, 80 OHIO ST. L.J. 399 (2019) (discussing IoT in the context of toys); Eldar Haber, *The Wiretapping of Things*, 53 U.C. DAVIS L. REV. 733 (2019) [hereinafter Haber, *The Wiretapping of Things*] (discussing IoT and wiretapping); Dan Feldman & Eldar Haber, *Measuring and Protecting Privacy in the Always-on Era*, 35 BERKELEY TECH. L. REV. 197 (2020) (suggesting computational solutions to privacy in IoT devices).

machine learning algorithms,[214] we are likely to advance to a stage where digital devices could sufficiently assess, at least to a high level of statistical probability, when a perilous situation occurs, often a crime being committed, and whether victims are involved.[215]

Consider the following running example. A woman runs her daily route while wearing a smart wearable, like a Fitbit or an Apple Watch.[216] After running several miles, the wearable detects that the user's vital signs show unusual levels of stress, faster heart beats than expected, and that her location indicates that she has strayed far off her regular, or any other runnable, running route. At this point, aggregation and analysis of the user's data could provide a statistical analysis of the likelihood that she might be in grave danger.[217]

---

214. Natural Language Processing (NLP) combined with Machine Learning (ML) "helps computers to autonomously learn tasks such as the recognition, understanding and generation of natural language (i.e. the language spoken by humans)." Thomas Margoni, *Artificial Intelligence, Machine Learning and EU Copyright Law: Who Owns AI?* 2 (CREATE, Working Paper No. 2018/12, 2018).

215. The algorithms that power many of current AI-based devices generally "use models of probability to make educated guesses." Matt Day et al., *Thousands of Amazon Workers Listen to Alexa Users' Conversations*, TIME (Apr. 11, 2019, 2:04 PM), https://perma.cc/WK8J-SXVU. It follows that AI could also make educated guesses on the probability of a crime being committed.

216. Wearable IoT devices are often used to monitor the body. *See* Erika J. Nash, *Notice and Consent: A Healthy Balance Between Privacy and Innovation for Wearables*, 33 BYU J. PUB. L. 197, 199 (2018) (listing various ways that wearables measure information about their users). Key examples in the fitness industry include Fitbit, Jump, and the Samsung and Apple smartwatches. These devices will often gather large quantities of data on the user's physical activities and other related health metrics, e.g., an individual's heart rate or quality of sleep. *See* ANDREW HILTS ET AL., EVERY STEP YOU FAKE: A COMPARATIVE ANALYSIS OF FITNESS TRACKER PRIVACY AND SECURITY 5–7 (2016), https://perma.cc/8D5D-28SK (PDF) (explaining how fitness wearables function and demonstrating the growth of sales by leaders in the fitness-wearable industry); Hillary Brill & Scott Jones, *Little Things and Big Challenges: Information Privacy and the Internet of Things*, 66 AM. U. L. REV. 1183, 1190–91 (2017) (citing the fitness industry as one of the primary drivers of wearable IoT devices).

217. Essentially, equipped with proper sensors, IoT devices could gather various types of data and when aggregated and analyzed, could indicate, to

Another example could be that of computerized personal assistants, like Amazon Echo or Google Home,[218] that detect, to a high probability, that a felony is currently being committed within one's house. Consider an Amazon Echo device that, while operating in one's house, detects aggregated signs of harsh domestic violence. To take this example further, suppose that using biometric voice recognition and identification, the Echo device computes that the domestic violence is committed by

---

some statistical probability at least, that a crime is being committed. These sensors could include, *inter alia*, emotion sensing (detecting, e.g., heartbeat and body temperature), emotion recording, and geolocation. *See, e.g.*, Jeong-Yong Bryun et al., *Internet of Things for Smart Crime Detection*, 7 CONTEMP. ENG'G SCIS. 749, 752 (2014) (proposing a method of crime detection relying on wearable sensors).

218. Computerized personal assistants—IoT devices—are often voice-activated and awaiting a voice command. Users can communicate with them for various purposes. *See* Ricky Philip, *Is the Future of Web Application Development Affected by the Disruptive Growth and Impact of IoT?*, IT CHRONICLES (Sept. 21, 2020), https://perma.cc/AV9R-TZMZ ("Personal assistants operate via an IoT database as they extrapolate users' requests, notes, and even answer their questions."). Google, for instance, described their personal assistant (Google Home) by "Ask your Google Assistant questions. Tell it to do things. It's your own Google, always ready to help." *See Google Nest Devices*, GOOGLE (2020), https://perma.cc/B338-UYSQ (describing different Google Home devices). Amazon released the Echo device, describing it as:

> [A] hands-free speaker you control with your voice. Echo connects to the Alexa Voice Service to play music, make calls, send and receive messages, provide information, news, sports scores, weather, and more—instantly. All you have to do is ask. Echo has seven microphones and beam forming technology so it can hear you from across the room—even while music is playing. Echo is also an expertly tuned speaker that can fill any room with 360° immersive sound. When you want to use Echo, just say the wake word "Alexa" and Echo responds instantly. If you have more than one Echo or Echo Dot, Alexa responds intelligently from the Echo you're closest to with ESP (Echo Spatial Perception).

Amazon Echo, AMAZON, https://perma.cc/3NMM-Y3N2; *see also* Haber, *The Wiretapping of Things*, *supra* note 213, at 745–47 (describing IoT devices, while differentiating between always-ready and always on devices). For more on computerized personal assistants, see *Top 22 Intelligent Personal Assistants or Automated Personal Assistants*, PREDICTIVE ANALYTICS TODAY, https://perma.cc/SM3D-ZLTH (analyzing and ranking personal assistant devices on the market).

members of the household, ruling out potential false positives that might result from unrelated background sounds.

The normative question that these two examples raise is, supposing that technology could provide rather accurate estimations of crimes (or other perilous situations) in real time, should these devices, or their service providers, be placed under a legal obligation to aid those in peril, and mostly those who fall victim to violent criminal activities, thereby notifying law enforcement agents? In other words, should policymakers extend the scope of Samaritan laws to include AI platforms and services?

Notably, while this discussion might seem in the realm of science fiction, it is not entirely so. For example, in one instance Amazon Echo allegedly notified police of an ongoing violent assault.[219] Amazon, however, has rejected these claims for now,[220] thus it is difficult to evaluate the existence of such practices by companies. What Amazon did confirm was that its employees, working to help improve the abilities of digital assistants and customer experience, were routinely listening to voice recordings captured by these devices and have heard recordings that were believed to be criminal in nature.[221] Amazon, however, instructed its workers not to report these incidents, under their legal interpretation that it is not Amazon's responsibility to interfere.[222] This last example,

---

219. *See* Christopher Mele, *Did an Echo Call 911 During a Domestic Assault? Amazon Says No*, N.Y. TIMES (July 11, 2017), https://perma.cc/ZMN7-77TG.

220. *See id*. (reporting the statement of an Amazon spokeswoman, Rachel Hass, who stated that for the device to have made a call, "the receiving end [that is, the police] [need[ed] to also have Alexa calling and messaging set up" and indicating that the police did not have this capability).

221. *See* Day et al.*, supra* note 215.

222. *See id*. Notably, in this instance, the workers were unable to identify which user made the recordings, thus, even if one would suggest that Samaritan duties could be imposed on the human element in this loop, and apart from the fact that these crimes were not livestreamed and have potentially ended, excludes them from any such Samaritan duty. *See id*. ("Employees do not have direct access to information that can identify the person or account as part of this workflow."). Apple and Google also hire teams to examine the interpretation of their personal assistants (Siri and Google Home, respectively), but removes personally identifiable information from

however, does not shed light on potential artificial Samaritans, as it merely echoes the online Samaritans paradigm.[223] But even if current AI technology might still not be placed in a proper position to serve as Samaritans, there are some indications that technology might be headed towards this direction.

Prior to normatively evaluating whether it is socially and legally desirable to impose affirmative duties on artificial bystanders, a few caveats should be set. First, it is important to note that this Part excludes the rather controversial debate, often categorized under the rubric of AI and ethics, on whether AI should have legal rights and, if so, to what extent.[224] While such debate holds scholarly importance, it captures disagreement that would not complement the present discussion on digital Samaritans. The second caveat relates to the potential overoptimistic nature of many to attribute AI with abilities it simply does not yet possess and perhaps never will.[225]

---

these recordings, thus their employees are unable to act upon it, even if purely criminal in nature. *See id.*

223. *See supra* Part III.A.2.

224. One of the prominent questions in respect to legal rights is whether machines deserve to have a right of free speech. For more on this discussion, see Stuart Minor Benjamin, *Algorithms and Speech*, 161 U. PA. L. REV. 1445, 1447 (2013) ("[I]f we accept Supreme Court jurisprudence, the First Amendment encompasses a great swath of algorithm-based decision—specifically, algorithm-based outputs that entail a substantive communication."); James Grimmelmann, *Speech Engines*, 98 MINN. L. REV. 868, 917–31 (2014) (assessing various theories of whether search engine rankings are speech protected by the First Amendment and concluding that such rankings are "descriptive opinions" of relevance which may only be actionable in tort when "subjectively dishonest"); Toni M. Massaro & Helen Norton, *Siri-ously?: Free Speech Rights and Artificial Intelligence*, 110 NW. U. L. REV. 1169, 1172 (2016) (examining whether computers with "strong AI"—AI that produces actual, independent thought—could hypothetically be treated as a speaker under the First Amendment); Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495, 1526 (2013) (arguing that the "functional nature of search engines" does not preclude application of the First Amendment to search results); Hugh McLachlan, *Ethics of AI: Should Sentient Robots have the Same Rights as Humans?*, INDEPENDENT (June 26, 2019, 13:49), https://perma.cc/2EHR-SEPV ("To deny conscious persons moral respect and consideration on the grounds that they had artificial rather than natural bodies would seem to be arbitrary and whimsical.").

225. For more on past overly optimistic predictions regarding the abilities of AI, see *What Should we Learn from Past AI Forecasts?*, OPEN PHILANTHROPY

Under this caveat, one might argue that the following debate on AI Samaritans is rather theoretical at this point, and might remain as such. More specifically, AI technology has not yet reached the stage at which it could accurately estimate the probability that a crime is being committed, and it might never reach it.

Can AI technology accurately report on crimes being committed? At the current stage of technology, it would be presumptuous to answer this question in the affirmative.[226] Truly, technology is advancing at a rapid pace, and it enables companies to capture and store mass amount of data on individuals, sometimes at any given time.[227] Wearable IoTs could monitor and store various types of data that could include heart rate, sexual activity, sleep patterns, steps taken, and geolocation, to name but a few examples.[228] It is expected that IoT and AI technology will continue to develop in years to come

(May 2016), https://perma.cc/SN3C-2YAD (last updated Sept. 2016) (summarizing the optimism during the 1950s through the 1980s regarding rapid development of AI).

226. *See infra* notes 232–236 and accompanying text.

227. I have termed this technology elsewhere as "always-on" devices, "equipped with sensors [that] could theoretically capture all data in the vicinity of these sensors, depending on the type of data that the sensor could capture." *See* Haber, *The Wiretapping of Things*, *supra* note 213, at 748–52 (explaining and offering examples of the "datafication of things"). For more on always-on devices and the "datafication of things," see Katharine Saphner, Note, *You Should Be Free to Talk the Talk and Walk the Walk: Applying* Riley v. California *to Smart Activity Trackers*, 100 MINN. L. REV. 1689, 1689–93, 1715 (2016).

228. To exemplify, Fitbit currently collects "identifiers," "demographic information," "commercial information," "biometric information," "internet or other electronic network activity information," "geolocation data," "electronic, visual, or similar information," "professional or employment related information," information provided by the user, and "inferences drawn from any of the above." *Fitbit Privacy Policy*, FITBIT (Dec. 18, 2019), https://perma.cc/QJD2-Y4EG*; see* Haber, *The Wiretapping of Things*, *supra* note 213, at 749 ("Wearable IoT devices or smart activity trackers could monitor various types of data, such as heart rate, sexual activity, sleep patterns, steps taken, calorie consumption, and geolocation.").

and will open a world of possibilities not necessarily imaginable right now.[229]

When it comes to law enforcement, AI has already started assuming a role in which the data gathered by these devices is often used *ex post* as evidence.[230] Currently, however, this technology might not be able to accurately detect when individuals are in a perilous situation. Even devices that are considered relatively innovative and in advanced stages of development might create many false positives or false negatives.[231] Computerized personal assistants might, for instance, capture someone saying, "You are killing me!" and determine that a crime is being committed. In practice, however, the exclamation might have been merely a euphemism, a joke, storytelling, rehearsal of a play, or simply sounds from the TV, to name a few examples.[232]

Some of these practical challenges might be met with the use of other technological advancements, like that of biometric data and voice recognition that could aid the device in identifying the speakers, thereby reducing chances that the data was not communicated by another person or medium.[233] However, even utilizing those assistive technologies, it would still be difficult for AI devices to evaluate and interpret the

229. *See, e.g.*, Louis Columbus, *2018 Roundup of Internet of Things Forecasts and Market Estimates*, FORBES (Dec. 13, 2018, 11:49 AM), https://perma.cc/HD49-HF7H ("Combined, businesses, governments, and consumers will invest nearly $1.6 trillion to install IoT solutions in 2020.").

230. *See, e.g.*, Christine Hauser, *Police Use Fitbit Data to Charge 90-Year-Old Man in Stepdaughter's Killing*, N.Y. TIMES (Oct. 3, 2018), https://perma.cc/A723-MJDR (describing how data from a Fitbit was used as evidence to resolve a murder); Hauser, *supra* note 6 (same).

231. *See infra* notes 232, 234–236 and accompanying text.

232. For various reported stories on how Amazon's Alexa misunderstood requests, or picked up requests from the TV, see Gia Liu, *Hey, I didn't order this Dollhouse! 6 Hilarious Alexa Mishaps*, DIGIT. TRENDS (Mar. 6, 2018), https://perma.cc/TC3Y-EBXA.

233. *See generally* Oscar Knagg, *Building a Speaker Identification System from Scratch with Deep Learning*, MEDIUM (Oct. 2, 2018), https://perma.cc/9MA8-FTHU (describing how to create a high-accuracy voice recognition system using neural networks).

context of human speech.[234] Here, a human witnessing such interaction is likely to comprehend the situation better than any machine.[235] Thus, to date, the available technology is probably not yet in a phase where it could compute human interactions better than humans, and, as mentioned, it could lead to many inaccurate computations.[236]

But as technology is quickly evolving, along with the financial investments in the field of AI and IoT, such crime notification could be possible, and perhaps will become a reality soon. Thus, assuming for now that AI technology might pass these thresholds and barriers and will be able to properly detect crimes in real time, the normative question is should the AI device—or the companies that analyze the users' data and reach a conclusion that a crime (or another perilous situation) is being committed—be legally obligated to contact law enforcement agencies, thus creating a new category of bystanderism in the form of "artificial Samaritans"? What are the benefits and drawbacks of imposing legal liability on technological

234. *See* Day et al., *supra* note 215 (articulating that AI devices have difficulties interpreting speech, "especially when grappling with new slang, regional colloquialisms or languages other than English").

235. Self-driving cars, for instance, have experienced problems in interpreting risks and hazards. One example is an Uber self-driving car, that "interpreted a pedestrian and her bike as a plastic bag or piece of cardboard," along with other mistakes that humans will not likely make. *See* Aarian Marshall, *The Uber Crash Won't Be the Last Shocking Self-Driving Death*, WIRED (Mar. 31, 2018, 07:00 AM), https://perma.cc/M4TG-7592 (emphasizing that "[e]ven little things"—such as small pieces of tape on road signs, shimming exhaust, and adverse weather conditions including fog—"have been observed to fool" AI systems).

236. Computerized personal assistants, those that are AI-based, often mistakenly interpret humans. Amazon Alexa, for instance, accidentally recorded a conversation of a husband and wife and sent it to one of their contacts without permission. *See* Hamza Shaban, *An Amazon Echo Recorded a Family's Conversation, Then Sent It to a Random Person In Their Contacts, Report Says*, WASH. POST (May 24, 2018, 6:40 PM), https://perma.cc/JD9D-R69E ("[T]he Echo woke up when it heard a word that sounded like 'Alexa' and interpreted any subsequent conversation as a 'send message' request"); *see also* Day et al., *supra* note 215 ("I think we've been conditioned to the [assumption] that these machines are just doing magic machine learning. But the fact is there is still manual processing involved." (alteration in original) (quoting Florian Schaub)).

companies in this context? And will it be legally plausible and normatively desirable?

Evaluating whether to impose obligations on AI operators to report crimes could first be linked to the theoretical roots and justifications of imposing Samaritan duties to begin with.[237] As mentioned in Part II, tracing its historical roots, duty-to-report statutes were often responsive to specific incidents that sparked a public outcry following humans' non-intervention in instances where they morally should have aided the victim or, at the very least, reported the crime. Hence, at least one of the rationales behind Samaritan laws is linked to the public shaping of what constitutes a socially acceptable or moral behavior.[238]

At this point, one might argue that the rationale should not extend to non-humans. These machines are not part of the social fabric.[239] They are also in no need of moral education.[240] On this last point, one might even argue that Samaritan laws are driven by the need to educate others on aiding—and less on the crime committed—i.e., that lacking a human to educate, there is little (if any) social need for "artificial Samaritan" duties. On the other hand, the fact that Samaritan laws emerged from a specific social need does not mean that future legislation or other forms of regulations should not tilt these moral rationales towards the aid to those in peril, or even, the gathering of crucial evidence.

Imposing artificial Samaritan duties on AI devices or service operators necessitates a social and legal cost-benefit analysis, translated here into potential benefits and drawbacks.

---

237. *See supra* notes 56–60 and accompanying text.

238. *See* Levy, *supra* note 41, at 617 (advocating for states to impose Samaritan laws on the basis of "formal recognition of a moral duty that we all owe to each other, a duty to attempt to save one another when the burden and risk are low and the potential benefits . . . are very high").

239. *See generally* Joanna J. Bryson, *The Past Decade and Future of AI's Impact on Society*, *in* TOWARDS A NEW ENLIGHTENMENT? A TRANSCENDENT DECADE 127 (11th ed. 2019) (presenting AI devices as powerful *tools* for humans to use, but which come with an inherent risk of devaluing human life and work in the future).

240. *See* Bryson, *supra* note 239, at 129–32 (defining artificial intelligence as a product of human coding and machine learning that allows the machine to (1) "perceive contexts for action"; (2) "act"; and (3) "associate contexts to actions").

Obviously, imposing Samaritan duties on AI operators could aid in crime prevention and detection in both public and private spheres.[241] When more technologies are readily available within public spheres, AI devices will be able to increase the personal safety of individuals, as these devices could contact law enforcement agencies when people are in danger and equip those agencies with crucial data to aid in ceasing the criminal activity or other perilous situations, while also obtaining evidence.[242] This argument might become even more important for those crimes that often go under the radar or where evidence is scarce. The mitigation of these crimes rely mostly on human reporting and these crimes often go undetected due to the location of the crime, i.e., within one's private sphere, like one's home.[243] Thus, imposing Samaritan duties could aid in the enforcement of criminal activities that are often not reported by victims and benefit victims who are unable to contact law enforcement agencies during the criminal conduct or otherwise perilous situation.

Moreover, imposing Samaritan duties on AI operators does not raise the major libertarian concern which Samaritan

241. *See supra* note 230 and accompanying text.

242. *See* Douglas A. Fretty, *Face-Recognition Surveillance: A Moment of Truth for Fourth Amendment Rights in Public Places*, 16 VA. J.L. & TECH. 430, 447 (2011) ("[B]y stepping in front of a face-identifying camera, a civilian is matched not only with his state-owned photograph but also any data associated with his name—residence, welfare status, employment, social security number, tax history, criminal record, child support compliance, etcetera."); Jeong-Yong Byun & Aziz Nasridinov, *supra* note 217, at 752–53 (enumerating a framework for AI detection of criminal activity).

243. Currently, many crimes go undetected simply because of the nature of the attack, and perhaps most commonly, crimes linked with domestic violence or sexual abuse. *See* Kaufman, *supra* note 19, at 1325 n.33 (quoting crime statistics indicating that "[o]nly 230 out of every 1,000 sexual assaults are reported to police"). To name two main reasons for why sexual violence victims are reluctant to report these crimes: they fear retaliation or have little faith that law enforcement agencies will aid them. *See id.* at 1337 (offering objections by survivors of sexual violence to mandatory reporting of sexual assaults). In addition, some individuals do not report or assist victims of crimes because they lack self-confidence to intervene. *See* Dickerson, *supra* note 165, at 62 (making the same argument with reference to bullying).

laws often raise—that of liberty.[244] Samaritan duties are often criticized as undesirable due to undermining personal autonomy—something that machines do not currently (and are not expected to soon) have.[245] Lacking any meaningful human element in the loop, Samaritan obligations will not impose direct liberty-limiting duties on humans, but rather, mostly on machines or, one might argue, AI companies.[246] In other words, if technology could aid in preventing, detecting, and even obtaining evidence of criminal activities or other perilous situations, then artificial Samaritan duties will presumably be beneficial for society as a whole.

At the same time, imposing a legal duty on AI operators and service providers to report crimes is both undesirable and implausible for various reasons, along with other potential challenges that arise from both legal and pragmatic aspects. While the devices might aid in increasing personal safety for

244. *See* Rogers, *supra* note 25, at 904–05 (examining the libertarian argument that Samaritan laws infringe on an individual's "right not to rescue").

245. *See* Montana, *supra* note 39, at 549 ("Opponents to duty to aid laws would argue that the personal autonomy principle should take precedent, especially when such laws aim to punish nonfeasance, not misfeasance—the key to liability for negligent and criminal acts in United States law."); Rogers, *supra* note 25, at 904–05 (analyzing the libertarian argument that duty-to-rescue laws "infringe upon individual freedom by denying people the choice of whether to assist a person in peril").

246. Truly, there are many humans involved in the process of creating and maintaining AI technology, thus they are not entirely exempt from the loop. *See, e.g.*, Vyacheslav Polonski & Jane Zavalishina, *Can We Build the Good Samaritan AI? Three Guidelines for Teaching Morality to Machines*, CTR. FOR PUB. IMPACT (Dec. 12, 2017), https://perma.cc/B8BV-63WN (emphasizing the *human* responsibility to teach machines to maximize fairness and to overcome racial and gender bias when making decisions). But within the context of live-reporting, other than having to program such ability within the device or service, the actual contacting will not be made by humans under this scenario. In other words, while requiring someone to program Samaritan duties could be considered as limiting the liberty of companies, it is not equivalent to the libertarian arguments in the context of human reporting. On the other hand, while AI Samaritan obligations will not impose direct liberty-limiting duties on humans, placing obligations on companies might in turn place limits on an individual's liberty to decide on what the machine is permitted to divulge to law enforcement agencies or others. Thus, there is an indirect restriction of liberty on some individuals, even when the direct duties are imposed on intermediaries.

some individuals,[247] mostly in specific events, such duties will trade off human rights and liberties in return.[248] Normatively and legally speaking, Samaritan duties on AI platforms could be considered as too intrusive from a human rights perspective, potentially infringing upon privacy and freedom of expression rights of individuals. Imposing such duties on AI companies could also increase fears of data misuse[249] by both private companies and, perhaps more dramatically, law enforcement and other intelligence agencies which will be able to receive more data on individuals without judicial safeguards or reliance on the companies' willingness to voluntarily disclose such information.[250] While criminal law always considers the impact on human rights and liberties when imposing any affirmative duty, the impact on these rights and liberties within the current Samaritan laws are inherently different within the context of AI, and thus must be rebalanced if imposed.

Within the context of human rights and liberties, the privacy challenge is perhaps the most dominant factor against imposing any artificial Samaritan duties. The privacy challenge begins with a constitutional debate.[251] The Fourth Amendment—the most relevant constitutional protection for

247. *See supra* note 243 and accompanying text.

248. *See* Haber, *The Wiretapping of Things*, *supra* note 213, at 792 (expressing that increased dependence on AI devices in one's daily life—i.e. the "smartification of things"—could "increase the potential threat to privacy and other civil rights and liberties due to datamining and data analysis capabilities").

249. *See* Andrew Smith, *Using Artificial Intelligence and Algorithms*, FED. TRADE COMM'N: BUS. BLOG (Apr. 8, 2020, 9:58 AM), https://perma.cc/S895-M2M5 ("The use of AI technology . . . to make predictions, recommendations, or decisions . . . presents risks, such as the potential for unfair or discriminatory outcomes or the perpetuation of existing socioeconomic disparities.").

250. *See* Bryson, *supra* note 239, at 139 (elucidating that AI enables third parties to access records on any individual that produces storable data and emphasizing that "[w]e are to some extent all celebrities now: any one of us can be identified by strangers, whether by facial-recognition software or datamining of shopping or social media habits").

251. *See* Daniel J. Solove, *Digital Dossiers and the Dissipation of Fourth Amendment Privacy*, 75 S. CAL. L. REV. 1083, 1085 (2002) ("The first source for protecting privacy against infringement by law enforcement agencies is the Fourth Amendment . . . .").

individuals' privacy—protects "persons, houses, papers, and effects" from unreasonable and warrantless searches and seizures.[252] This constitutional protection should extend generally to the data that is gathered from AI devices, as they will likely be categorized as "effects."[253] Even if not, many of these AI devices will likely be present within an individual's house or worn by him or her and thus could generally fall under the Fourth Amendment protection of either houses or persons, respectively.[254] Under the Supreme Court's construction of the so-called reasonable expectation of privacy test, a Fourth Amendment "search" is conducted when there is both a subjective and objective expectation of privacy by individuals,[255]

252. *See* U.S. CONST. amend. IV; Katz v. United States, 389 U.S. 347, 359 (1967) (emphasizing that the Fourth Amendment protection against unreasonable searches and seizures extends to "[w]herever a man may be"). For more on the Fourth Amendment, see generally Solove, *supra* note 251. Notably, along with various protections by federal and state legislation, the right to privacy had been interpreted to be protected under the Supreme Court's interpretation of the Bill of Rights other than the Fourth Amendment, mostly within the First, Third, Fifth, and Fourteenth Amendments. *See* U.S. CONST. amends. I, III, V, XIV; *Katz*, 389 U.S. at 350 n.5 (referencing the ways in which the First, Third, and Fifth Amendments protect personal privacy); Griswold v. Connecticut, 381 U.S. 479, 500 (1965) (Harlan, J., concurring) (asserting that a right to privacy is incorporated to the states through the 14th Amendment Due Process Clause).

253. U.S. CONST. amend. IV; *see* Haber, *The Wiretapping of Things*, *supra* note 213, at 752 (making this argument as such technologies constitute "an individual's personal property"); Andrew Guthrie Ferguson, *The Internet of Things and the Fourth Amendment of Effects*, 104 CALIF. L. REV. 805, 853–64 (2016) (arguing that an interpretation of IoT devices as Fourth Amendment "effects" is consistent with Fourth Amendment theory and interpretations of protection for persons, houses, and papers).

254. Compare Justice Potter Stewart's view in the seminal *Katz v. United States* opinion, wherein he articulates that "[t]he Fourth Amendment protects people, not places." *Katz*, 389 U.S. at 351.

255. *See id.* at 359 (forming the reasonable expectation of privacy test, which has a twofold requirement to determine whether governmental conduct constitutes a search); Riley v. California, 573 U.S. 373, 381–82 (2014) ("[T]he ultimate touchstone of the Fourth Amendment is 'reasonableness.'" (quoting Brigham City v. Stuart, 547 U.S. 398, 403 (2006))); *see also* William Baude & James Y. Stern, *The Positive Law Model of the Fourth Amendment*, 129 HARV. L. REV. 1821, 1829 (2016) (articulating the two-part test as (1) determining whether a government agent committed an act "that can be characterized as either a search or seizure;" then (2) evaluating the act—if either a search or

that is, unless there is an exception to general warrant requirement.[256]

But interpretation of current Fourth Amendment jurisprudence will likely lead to the outcome that the Fourth Amendment will not apply in most of these instances, as long as data sharing is consensual,[257] and as long as the data is not stored only locally on the device. Under the so-called

---

seizure—for its reasonableness); Jonathan Mayer, *Government Hacking*, 127 YALE L.J. 570, 581–82 (2017) (reciting the reasonableness component of the test as judicial scrutiny of "the privacy procedures associated with" a government "search").

256. Courts have created exceptions to the general reasonable expectation of privacy test. These examples include, *inter alia*, the Terry stop and frisk search (requiring reasonable suspicion rather than probable cause), items displayed in plain view during a search, exigent circumstances, and consensual searches. *See* Maryland v. Macon, 472 U.S. 463, 469–70 (1985) (concluding that an undercover officer did not conduct a Fourth Amendment search or seizure when he, lacking a warrant, viewed and purchased adult magazines that were "intentionally exposed" on a store's shelves, then arrested the store clerk for distribution of obscene material); Coolidge v. New Hampshire, 403 U.S. 443, 455, 466 (1971) (emphasizing that to overcome the presumption that extra-judicial searches are *per se* unreasonable under the Fourth Amendment, "there must be a showing by those who seek exemption . . . that the exigencies of the situation made that course imperative" (internal quotations omitted)); Terry v. Ohio, 392 U.S. 1, 30–31 (1968) (holding that an officer does not violate the Fourth Amendment for conducting a frisk search if he (1) has a reasonable suspicion that the person has committed, is committing, or is about to commit a crime and (2) has a reasonable belief that the person may be armed and presently dangerous); *see generally* Benjamin T. Clark, *Why the Airport and Courthouse Exceptions to the Search Warrant Requirement Should Be Extended to Sporting Events*, 40 VAL. U. L. REV. 707, 715–23 (2006) (analyzing the consent, *Terry* stop and frisk, airport, and courthouse exceptions to the warrant requirement for a reasonable search and applying those exceptions to justify warrantless searches at sporting events); Haber, *The Wiretapping of Things*, *supra* note 213, at 753 (noting that barring exceptions, law enforcement agencies "are required to obtain a warrant in most instances" before conducting a search).

257. Courts have repeatedly held that a consensual access to one's home, even if predicated upon mistrust, nullifies Fourth Amendment protection. *See* Hoffa v. United States, 385 U.S. 293, 302 (1966) ("Neither this Court nor any member of it has ever expressed the view that the Fourth Amendment protects a wrongdoer's misplaced belief that a person to whom he voluntarily confides his wrongdoing will not reveal it."); Katherine J. Strandburg, *Home, Home on the Web and Other Fourth Amendment Implications of Technosocial Change*, 70 MD. L. REV. 614, 652–53 (2011) ("No court has held that the Constitution protects against misplaced trust.").

"third-party doctrine," and along with some notable exceptions,[258] individuals who share information with a third party who is the intended recipient of the information—the AI platform in this instance—have no reasonable expectation of privacy in that data.[259] In other words, when individuals communicate with IoT devices and know that the gathered data could theoretically be subjected to Samaritan duties, these AI operators are permitted to legally divulge such information to law enforcement agencies without raising potential constitutional violations.[260]

---

258. At least in some jurisdictions, the contents of emails held by internet service providers will not be subjected to the third-party doctrine. *See* United States v. Warshak, 631 F.3d 266, 285–86 (6th Cir. 2010) ("Given the fundamental similarities between email and traditional forms of communication, it would defy common sense to afford emails lesser Fourth Amendment protection."); Haber, *The Wiretapping of Things*, *supra* note 213, at 756. ("[A]t least some types of communication will receive Fourth Amendment protection not subjected to the third-party doctrine[,] . . . includ[ing] the contents of emails . . . ."). For other exceptions see, for example, United States v. Jacobsen, 466 U.S. 109, 114 (1984) ("Letters and other sealed packages are in the general class of effects in which the public at large has a legitimate expectation of privacy . . . ."); *In re* Search of Info. Associated with the Facebook Acct. Identified by the Username Aaron.Alexis that is Stored at Premises Controlled by Facebook, Inc., 21 F. Supp. 3d 1, 6–7 (D.D.C. 2013) (rejecting a search warrant application calling for account information for a Facebook user, stating that "it unduly invaded the privacy of third parties" because the government would see "irrelevant" communications "sent by persons who could not possibly have anticipated that the government would see what they posted"); Carpenter v. United States, 138 S. Ct. 2206, 2217 (2018) (stating that accessing historical records containing the physical locations of cellphones is subject to Fourth Amendment protection).

259. Two leading Supreme Court cases constructed the third-party doctrine. *See* United States v. Miller, 425 U.S. 435, 443 (1976) ("[T]he Fourth Amendment does not prohibit the obtaining of information revealed to a third party . . . even if the information is revealed on the assumption that it will be used only for a limited purpose and the confidence placed in the third party will not be betrayed."); Smith v. Maryland, 442 U.S. 735, 744 (1979) (emphasizing that when one reveals information to a third party, he "assume[s] the risk" that the information may be revealed to authorities). For more on the third-party doctrine, see generally Stephen E. Henderson, *Beyond the (Current) Fourth Amendment: Protecting Third-Party Information, Third Parties, and the Rest of Us Too*, 34 PEPP. L. REV. 975, (2007); Orin S. Kerr, *The Case for the Third-Party Doctrine*, 107 MICH. L. REV. 561 (2009).

260. *See* cases cited *supra* note 259.

The privacy question further extends far beyond the constitutional level. There could be many forms of legislation or regulation that will restrict governmental access to the data in general, like that of wiretapping devices or using other legal frameworks to obtain stored data, but they might merely mirror the rationale to protect privacy in some instances.[261] The fear here is that imposing Samaritan duties on AI companies, if at all legally and practically possible, will undermine users' privacy in the same ways that the Constitution was set to protect privacy interests more than two centuries ago. It will broaden the ways law enforcement agencies can surveil individuals,[262] and such data collection and analysis might lead to misuse, thereby threatening democracy. It would impose substantial constraints on individuals' rights and liberties, and perhaps mostly on their right to privacy and on their right to free speech, as individuals will be reluctant to act freely in places where they should feel safe, such as in their own homes.[263] If we acknowledge that privacy protection, like that

261. A key example is that of the Wiretap Act, The Stored Communications Act, and the Pen Registers and Trap and Trace Devices Statute. *See* Omnibus Crime Control and Safe Streets Act of 1968, Pub. L. No. 90-351, 82 Stat. 197 (codified as amended at 18 U.S.C. §§ 2510–2522) (regulating real-time access to communication, including electronic communication); Electronic Communications Privacy Act of 1986, Pub. L. No. 99-508, §§ 201–02, 100 Stat. 1848, 1860–1868 (codified as amended at 18 U.S.C. §§ 2701–2713) (regulating access to the content and metadata stored by electronic communications services); Electronic Communications Privacy Act of 1986, Pub. L. No. 99-508, § 301–02, 100 Stat. 1848, 1868–1872 (codified as amended at 18 U.S.C. §§ 3121–3127) (regulating access to devices that obtain information about calls); Haber, *The Wiretapping of Things*, *supra* note 213, at 740–41 (noting that, in response to the impact of new technological developments on individual privacy, Congress implemented the Electronic Communications Privacy Act to modify the Wiretap Act). For more on the stored communication act in the context of social media platforms, see Strandburg, *supra* note 257, at 643–48.

262. For more on governmental surveillance in the context of national security, see generally Niva Elkin-Koren & Eldar Haber, *Governance by Proxy: Cyber Challenges to Civil Liberties*, 82 BROOK. L. REV. 105 (2016).

263. The importance of protecting the home is often reflected within courts' interpretations of the protection granted by the Fourth Amendment. *See, e.g.*, Silverman v. United States, 365 U.S. 505, 511–12 (1961) (discussing the significance of an individual's home within the Fourth Amendment

afforded at home by the Fourth Amendment, is crucial for individuals' liberty,[264] autonomy,[265] and even democracy,[266] then when homes become virtual, it would be difficult to argue why there should be such a vast difference in the privacy protection of them.[267] Thus, privacy rights are closely linked in this scheme with other human rights and liberties, such as freedom of expression, and might be at great risk.

The pragmatic issues of "artificial Samaritans" are non-negligible as well. Say that the rationale of fighting crimes that often go undetected due to lack of human reporting or domestic crimes within someone's home is considered by policymakers as highly important.[268] Would those who often commit such crimes—and their victims who are reluctant to notify—have their houses or bodies equipped with such devices?[269] What will be the social response of users, even those

protection); Strandburg, *supra* note 257, at 650, 658 ("The idea that the home is deserving of particular protection against government intrusion is deeply embedded in jurisprudence, culture, and popular and legal intuition.").

264. *See* Charles Fried, *Privacy*, 77 YALE L.J. 475, 483 (1968) ("Most obviously, privacy in its dimension of control over information is an aspect of personal liberty."); Sonia K. McNeil, Note, *Privacy and the Modern Grid*, 25 HARV. J.L. & TECH. 199, 205–06 (2011) ("Privacy . . . is a facet of personal liberty, moral autonomy, and democracy.").

265. *See* McNeil, *supra* note 264, at 206 (emphasizing the role of privacy in attaining individual goals).

266. *See* Ruth Gavison, *Privacy and the Limits of Law*, 89 YALE L.J. 421, 455 (1980) ("Privacy is also essential to democratic government because it fosters and encourages the moral autonomy of the citizen, a central requirement of a democracy."); McNeil, *supra* note 264, at 205–06.

267. For more on the potential shift to social media as a primary form of interaction, see Strandburg, *supra* note 257, at 655–57.

268. *See, e.g.*, Proclamation No. 8769, 3 C.F.R. 8769 (2011) ("In our schools and in our neighborhoods, at home and in workplaces across our Nation, stalking endangers the physical and emotional well-being of millions of American men and women every year. Too often, stalking goes unreported and unaddressed, and we must take action against this unacceptable abuse.").

269. *See, e.g.*, Kaufman *supra* note 18, at 1325 n.33 (presenting statistics of sexual assault victims who do not report their assault); Dickerson, *supra* note 165, at 61 ("Some victims [of bullying] simply suffer in silence because they fear ostracism, retaliation, or escalation in the bullying.").

who claim that they "have nothing to hide"[270] from law enforcement agencies but fear the misuse of such power? There is a valid chance that such legal duties will drive many users away from using AI technology, hence creating a chilling effect on its use and on innovation.[271]

Moreover, while for humans non-reporting of crimes or other perilous situations might be considered a morally wrong behavior, it does not currently apply to machines. The question is whether the notion of morality can be extended to those that operate and designed the technology. In other words, is it appropriate to impose moral obligations on those AI companies that produce the product or provide its related services? As Part II.A argued, imposing such moral obligations will reshape the liability safeguards that federal law seeks to provide communication technologies under § 230 of the Communication Decency Act.[272] While § 230 deserves much criticism on various grounds,[273] its reconstruction must be carefully evaluated due to the negative effects it might have on markets and innovation.

It is also important to consider the fears that AI platforms will transform into law enforcement agents and serve as a proxy for law enforcement—all without the constitutional or other legal safeguards set to protect individuals, which do not generally apply to these platforms as they are not state actors

---

270. The "nothing to hide" argument is mostly linked to the debate between privacy and security, whereas, arguably, individuals should not fear governmental surveillance programs that detect criminal activities, if they are not involved in criminal activities. *See* Daniel J. Solove, *"I've Got Nothing to Hide" and Other Misunderstandings of Privacy*, 44 SAN DIEGO L. REV. 745, 745–53 (2007) (explaining the "nothing to hide argument"). But these arguments are considered by many scholars as based on mistaken views of privacy. *See id.* at 764 ("[T]he problem with the nothing to hide argument is the underlying assumption that privacy is about hiding bad things."). For more on the "nothing to hide" argument, see generally DANIEL J. SOLOVE, NOTHING TO HIDE: THE FALSE TRADEOFF BETWEEN PRIVACY AND SECURITY 21–33 (2011).

271. For a thorough analysis of the privacy-innovation tradeoff that occurs in many instances, see generally Tal Z. Zarsky, *The Privacy-Innovation Conundrum*, 19 LEWIS & CLARK L. REV. 115 (2015).

272. *See* 47 U.S.C. § 230.

273. *See, e.g.*, Freiwald, *supra* note 192, at 594–96 (discussing criticism to the immunity provision of § 230).

per se.[274] If AI companies expand their role as proxies for law enforcement, then many individuals will likely be reluctant to use this technology, or even if they do use it, they will be very careful in what they say in its vicinity.[275] They might even fear false positives and thus refrain from any speech that could be interpreted by the AI device as potential criminal activity or a perilous situation.[276] To simplify this argument, while placing speech constraints might be considered acceptable in some contexts, such as not saying "bomb" in an airport, people should certainly feel that they can say "bomb" in their house.[277]

The cost-benefit analysis of imposing duties to report on AI devices currently tilts in the direction of not imposing artificial Samaritan duties. It will be proven as impractical (due to the current state of technology and legal constraints) and, perhaps more importantly, undesirable from both a legal and social perspective. At the same time, there should be little doubt

274. To be considered as a state actor, one must exercise powers that are "traditionally the exclusive prerogative of the State." Jackson v. Metro. Edison Co., 419 U.S. 345, 353 (1974). Thus, the First Amendment only applies to a state action, meaning that under current interpretation, online platforms are not considered as acting on behalf of the state. *See* Marsh v. Alabama, 326 U.S. 501, 508–09 (1946) (establishing the state actor doctrine); Blum v. Yaretsky, 457 U.S. 991, 1005 (1982) (applying the exclusive public function test); Langdon v. Google, 474 F. Supp. 2d 622, 631–32 (D. Del. 2007) (concluding that Google, Yahoo, and Microsoft are private companies and working with state universities did not make them state actors subject to the First Amendment); *see also* Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV 1598, 1609–13, 1658 (2018) (reviewing the state action doctrine in the context of assessing intermediary liability for content on online platforms).

275. *See* Chavie Lieber, *Amazon's Alexa Might be a Key Witness in a Murder Case*, VOX (Nov. 12, 2018, 5:00 PM), https://perma.cc/92SL-TX9M ("Americans are just waking up to the fact that their smart devices are going to snitch on them. And that they are going to reveal intimate details about their lives that they did not intend law enforcement to have." (quoting an unnamed privacy expert)).

276. *See* Mele, *supra* note 219 (describing an instance of Siri contacting authorities in response a news station airing the phrase "'Hey Siri, call 911'" which led to flood of accidental calls to emergency dispatchers from phones responding to the command).

277. *See* Jackie Salo, *'Stupid Joke' About Bomb on Plane Gets Woman Arrested*, N.Y. POST (Feb. 14, 2019, 8:17 PM), https://perma.cc/3NHN-VXB2 (stating that joking about a bomb in her carry-on bag led to the woman's arrest for making "a false report over a bomb threat").

that AI technology will play a role, and perhaps a substantial one, in law enforcement in the years to come.[278] Beyond the current expanding use of the data gathered by AI or IoT devices for enforcement purposes, both *ex post* and in real time,[279] it is likely that the smartification of the public sphere with the combination of biometric identification will become at least part of the enforcement matrix.[280] But as important as personal and public safety might be, harnessing AI technology for enforcement purposes must be carefully tailored to consider the negative impact on human rights and liberties.[281]

Overall, Samaritan duties are not currently likely to legally expand to the digital era in liberal and democratic societies anytime soon,[282] at least not without proper safeguards for civil rights and liberties.[283] The duty imposed in some jurisdictions to aid those in imminent risk in the physical world

278. *See* Jeong-Yong Byun & Aziz Nasridinov, *supra* note 217, at 752–53 (describing a proposed system for utilizing IoT devices to detect, visualize, and predict crime).

279. While it is still uncertain to what extent enforcement agencies are making use of IoT wiretaps, reports indicate that the Department of Justice considers activating cellphones' microphones as a legitimate practice. *See* Haber, *The Wiretapping of Things*, *supra* note 213, at 765–66; (highlighting the requirements that law enforcement agencies must meet to utilize wiretap devices and the increasing use of such devices); Declan McCullagh, *FBI Taps Cell Phone Mic as Eavesdropping Tool*, CNET (Dec. 4, 2006, 6:56 AM), https://perma.cc/K7VT-MJS3 (articulating the Department of Justice's approval for use of the roving bug wiretapping technique to eavesdrop on the infamous Genovese crime family).

280. *See generally* Fretty, *supra* note 242 (exploring the implications of law enforcement use of facial recognition on privacy)*; see also* Haber, *The Wiretapping of Things*, *supra* note 213, at 792–93 (discussing the potential for law enforcement to exploit smartification of things to combat crime).

281. *See* Haber, *The Wiretapping of Things*, *supra* note 213, at 792–93 (describing the need to establish checks and balances on law enforcement's ability to exploit personal data).

282. Notably, Corinne Moini argued that in the context of suspected child abuse or neglect, IoT companies that are covered by the Children's Online Privacy Protection Act (COPPA) should be obliged to report it to the state. *See generally* Corinne Moini, *Protecting Privacy in the Era of Smart Toys: Does Hello Barbie Have a Duty to Report*, 25 CATH. U. J.L. & TECH. 281 (2017).

283. *See id.* at 316 (discussing the privacy implications of duties to report abuse based on information heard by AI devices).

is highly contested to begin with,[284] and it will likely become much more controversial if new technologies enter this equation. It will also be undesirable to a great extent to expand such duties in the digital era, for both online and artificial bystanders.[285] While the reporting of both perilous situations and crimes committed with only AI awareness are generally important for safeguarding those in peril, and subsequently society as a whole, it is not likely that the law will play a substantial role in regulating such behavior, if any at all.[286]

But the fact that the current legal framework—crafted to shape human conduct in specific physical situations—is not aligned with the digital world, does not mean that the internet must figuratively become the Wild West or a no man's land in this respect. As the internet is likely to keep expanding into various domains of our lives,[287] the frequency of online bystanderism could equally rise, leaving the question of regulating proper human conduct to non-legal means. If society deems online Samaritans as important, then any proper solution must combine the potential benefits of using technological developments to better safeguard personal safety without placing constraints on free speech or violating individuals' privacy.[288] This might mean Samaritan duties for online reporting—whether formal or informal—will be left mostly under the discretion of users, platforms, and other indirect or informal policymaking.

While digital Samaritan duties are not likely to be regulated by the law, the discussion regarding both online and artificial bystanders is by no means unfruitful. It sheds much

284. *See supra* Part II.

285. *See supra* Part III.A.2; *see also supra* notes 248–249 and accompanying text.

286. *See supra* notes 248–249 and accompanying text.

287. *See, e.g.*, Katherine E. Tapp, Note, *Smart Devices Won't Be Smart until Society Demands an Expectation of Privacy*, 56 U. LOUISVILLE L. REV. 83, 89 (2017) ("[T]he [Internet of Things] embraces a reality where devices ranging from cell phones to wearables to our washing machines all connect to one another, building a huge, widely-varying network of connectivity.").

288. *See* Montana, *supra* note 39, at 556 (proposing an online Samaritan solution wherein "[t]he reporting itself can be done anonymously to dispel any fears over the bystander's privacy or personal wellbeing").

light on the role that the law often assumes in shaping moral conduct and how it might move into the hands of platforms that govern much of our daily lives.[289] In other words, the role of the law to shape the moral duties of society might become less effective due to the rise in the roles that both users and online platforms play in this regard, both likely to expand in the years to come.

IV. Samaritans and the Rise of Platform Governance

Online platforms are becoming integral to our current algorithmic society. Many people use these platforms to replace the once-physical everyday tasks, like shopping and communicating, with the digital arena.[290] There is little doubt that the way in which the architecture of these platforms is constructed could greatly influence human conduct,[291] as platforms are slowly becoming a new form of governors.[292] If the role of the state was once, *inter alia*, to convey and govern how norms and morality should be shaped in society, this role might partially be privatized by for-profit companies, potentially reshaping the ways social values are constructed.[293]

---

289. *See supra* note 58 and accompanying text; *see also infra* notes 292–296 and accompanying text.

290. *See, e.g.*, Zaryn Dentzel, *How the Internet Has Changed Everyday Life*, *in* CH@NGE: 19 KEY ESSAYS ON HOW THE INTERNET IS CHANGING OUR LIVES 235, 240 (6th ed. 2014) ("In almost everything we do, we use the Internet. Ordering a pizza, buying a television, sharing a moment with a friend, sending a picture over instant messaging.").

291. Lawrence Lessig argued that "code is law" and that architectures regulate behavior much prior to the current rise of platform governance. *See* LAWRENCE LESSIG, CODE: VERSION 2.0 120–37 (2006) (asserting that the internal architecture of cyberspace "constitute[s] a set of constraints on how [users] can behave," directly comparable to "real-space" regulations of human behavior such as the taxation of cigarettes to reduce human consumption and the enforcement of social norms).

292. *See* Klonick, *supra* note 274, at 1662–64 (conceptualizing modern platforms' moderation schemes, terms of use policies, role in democracy, and reflection of social norms as forms of governance).

293. *See, e.g.*, Laura DeNardis & Andrea Hackl, *Internet Governance by Social Media Platforms*, 39 TELECOMM. POL'Y 761, 766 (2015) (citing Twitter's temporary suspension of a journalist's account during the 2012 Olympics as demonstrative of "the power private companies have in determining the

In the context of Samaritan duties, the analogy of livestreaming crimes to real-time crimes could well serve as a vital example of how online platforms will shape or mirror the boundaries of morality in an algorithmic society.[294] Notably, however, online platforms do not operate in a vacuum.[295] Aside from potential state interference, online platforms rely on their users to convey their own attitudes towards the conduct of these for-profit platforms who strive to preserve their consumers and thus must be somewhat attentive to their desires.[296]

It might be preferable and efficient to have platforms and users negotiate the online playing field without legal interference, as it might be generally desirable from an economic and perhaps even a social perspective.[297] We might not need new laws and regulations to keep pace with technological changes; rather, we need users to actively respond to the shaping of moral obligations that are set by platforms.[298] Such self-regulation could aid in reframing the ways in which human

conditions of participation in the public sphere"); Klonick, *supra* note 274, at 1631–35 (describing the gradual development of content moderation guidelines from instructing moderators to remove content that produced a "gut" bad feeling to detailed manuals of what platforms perceive as socially and morally acceptable to display); Tarleton Gillespie, *Platforms are not Intermediaries*, 2 GEO. L. TECH. REV. 198, 199 (2018) ("Companies are beginning to actually grapple with how best to be stewards of public culture, a responsibility that was not evident to them at the start.").

294. *See supra* Parts III.A.1–2.

295. *See* Klonick, *supra* note 274, at 1662–64 (noting that online platform users both influence platforms' internal governance and self-regulate through enforcement of social norms).

296. *See id.* at 1663 ("[P]olicies and rules are modified and updated through external input; platforms are economically subject to normative influence of citizen-users . . . .").

297. Adam Smith has famously argued that market players acting in their own self-interest will react to consumers' demand—promoting the social good. *See generally* ADAM SMITH, AN INQUIRY INTO THE NATURE AND CAUSES OF THE WEALTH OF NATIONS (Sálvio Marcelo Soares ed., 4th ed. 2007).

298. As Jack Balkin acknowledged in the context of free speech, "[t]he problems of free speech in any era are shaped by the communications technology available for people to use and by the ways that people actually use that technology." Balkin, *Free Speech in the Algorithmic Society*, *supra* note 194, at 1151.

rights and liberties should be protected in this era,[299] which might better accommodate the technological and social changes we have experienced since the establishment of the modern state. This was in fact one of the main reasonings of the court in the seminal interpretation of § 230 in *Zeran v. America Online, Inc.*,[300] when granting broad immunity to internet service providers.[301] Aside from the need to encourage free speech and e-commerce, the court recognized the congressional purposes of encouraging "interactive computer services and users of such services to self-police the Internet for obscenity and other offensive material."[302]

Users are seemingly well-positioned to decide how they react to livestreaming crimes or other streams featuring perilous situations, as they could react to the platform's attitudes.[303] Some users might feel unobligated to personally

299. *See* M. Todd Henderson, *Why Self-Regulation of Social Media Could Work—The Financial Services Model*, THE HILL (July 29, 2019), https://perma.cc/7PSP-MN7K (explaining that it's incumbent on the internet industry to self-regulate in order to protect their own interest in a manner that aligns with the public's interest); Klonick, *supra* note 274, at 1630 (remarking that, whether rooted in corporate responsibility and identity or economic reasons, the development of platforms' content-moderation systems to reflect the normative expectations of users is precisely what the creation of the Good Samaritan provision in § 230 sought); Balkin, *Free Speech in the Algorithmic Society*, *supra* note 173, at 1209–10 (comparing the new set of social responsibilities faced by new media companies in the twenty-first century to the challenges faced by American journalists in the twentieth century).

300. 129 F.3d 327 (4th Cir. 1997).

301. *See id*. at 330 (noting that the purpose of intermediary immunity in § 230 was not only to protect the free speech of platform users but also to incentivize platforms to remove indecent content).

302. Batzel v. Smith, 333 F.3d 1018, 1028 (9th Cir. 2003) (citing 47 U.S.C. § 230(b)(4)); 141 CONG. REC. H8469–70 (daily ed. Aug. 4, 1995) (statements of Reps. Cox, Wyden, and Barton) (citing *Zeran*, 129 F.3d at 331 and Blumenthal v. Drudge, 992 F. Supp. 44, 52 (D.D.C. 1998)); *see* Klonick, *supra* note 274, at 1607–08 (explaining that the court in *Zeran* recognized two distinct congressional purposes for granting immunity under § 230: (1) as a good Samaritan provision to encourage interactive computer services and users to self-police the Internet and other offensive material, and (2) as a free speech protection for users).

303. *See* Danielle Keats Citron & Helen Norton, *Intermediaries and Hate Speech: Fostering Digital Citizenship for Our Information Age*, 91 B.U. L. REV. 1435, 1436 (2011) (advocating that internet intermediaries should adopt

report online criminal activities, even if they otherwise lack cognitive biases or failures, have no doubt in the authenticity or time of the event, or that their actions could potentially save people's lives.[304] Many others likely will. If the tables were turned, many users would have likely desired aid in a similar situation or simply consider the act of assisting those in peril as important.[305] Thus, these users could try nudging companies toward shaping their platforms to enable the desires of some to report livestreaming crimes (along with other perilous situations) or to even enable AI-based devices to become some form of guardians of personal safety.[306]

In translating these desires into viable solutions, what seemingly will matter most will rely on users' abilities to communicate their desires to platforms and other users—thus creating an incentive for platforms to invest and offer services that meet these desires. Under this assumption, users could demand, *inter alia*, that platforms at the very least enable quick and easy "SOS" calls within users' smart devices, by which users could simply shout a trigger phrase that will automatically contact law enforcement agencies or merely contact a trusted person and equip them with data on the perilous situation at hand.[307] Such a demand could be conveyed by users abstaining from using platforms that are not aligned with their moral

---

accessible and transparent policies that educate users about their rights and responsibilities as digital citizens by challenging hateful speech by responding with counter-speech and empowering community members to enforce norms of digital citizenship).

304. *See* Latané & Darley, *supra* note 166, at 244–49 (explaining the potential hinderance to bystander intervention posed by bystander effect, the term used to describe the concept that the presence of other bystanders might reduce the likelihood of bystanders to intervene, and the related diffusion of responsibility); Swan, *supra* note 34, at 984–86 (examining the bystander effect).

305. *See* WALLACE, *supra* note 34, at 198 (advancing the theory that the bystander effect might not be fully translated into the digital era, as there is evidence to support the conclusion that people tend to behave kindly toward strangers on the internet more so than in real-life situations).

306. *See* Balkin, *Free Speech in the Algorithmic Society*, *supra* note 194, at 1188 ("[Users] can use the digital public sphere to place social pressure on these digital platforms to modify their policies.").

307. *See infra* notes 313–315 and accompanying text.

perceptions;[308] communicating to other users (and the community as a whole) how platforms act (or misbehave), thus calling them out;[309] using traditional media to echo their message, which in turn could spark a public outcry;[310] or conveying their content and discontent in any other meaningful way. Platforms, in turn, might react to their users' desires.[311] They might reshape their contractual relationships with users, whether they are labeled as end-user license agreements, terms of service agreements, or community standards, striving for an optimal arena for users.[312]

Generally, we might witness a move of shifting Samaritan liabilities to self-reporting, and platforms will aid users to instantly report criminal activity or other forms of perilous situations, and thereby connect them to enforcement agencies. Interestingly, many companies have already made contacting emergency services readily accessible,[313] or have even provided

308. *See* David G. Post, *Anarchy, State and the Internet: An Essay on Law-Making in Cyberspace*, 1995 J. ONLINE L. 8 art. 3, para. 42 https://perma.cc/Y3S9-LURF (PDF) (claiming corporate competition between individual online platforms results in a "market for rules," which allows users to seek networks that have speech and conduct "rule sets" to their liking).

309. *See* Klonick, *supra* note 274, at 1652–53 (offering examples of platform policy change following collective action by users upset with platform policies).

310. *See id*. (remarking that while the media does not have a major role in changing platform policy per se, platforms have historically been responsive when media coverage is coupled with either the collective action of users or a public figure's involvement).

311. *See* Balkin, *Free Speech in the Algorithmic Society*, *supra* note 194, at 1201 ("[Platforms] champion a set of enlightened values that they believe that their end-users want—or should want—but they implement these values through bureaucracy and code without taking any sort of vote."); Klonick, *supra* note 274, at 1666 ("[W]hile it initially seems like a positive source of accountability that these systems are indirectly democratically responsive to users' norms, it also creates inherently undemocratic consequences.").

312. *See* Klonick, *supra* note 274, at 1648–49 (illustrating this concept by noting that some platforms, such as Facebook, are constantly updating their policies in large part to reflect the norms and expectations of its users).

313. For instance, many of Apple's iPhones and iWatches could be easily triggered to call emergency response services with a long touch of a button. Notably, making emergency SOS calls more accessible for users might in turn lead to many false positives, as the watch might accidently call these services

an emergency service in some instances, by which when the device detects danger it will ask the user whether to contact emergency services or not.[314] Regardless of the rationales behind such movements, some platforms have in fact enabled the use of various skills, like that of Alexa Guard or an intruder alert, that could respond to perceived intruders by making them believe that someone is present in the house or even make them believe that Alexa has contacted emergency services.[315]

Platforms have also directly reacted to livestreaming crimes.[316] Facebook, for instance, does not allow those who "proclaim a violent mission or are engaged in violence" to have an account, which includes, *inter alia*, those involved in organized violence or criminal activity (including terrorist

even without an emergency, often when the device is in repair and refurbishment facilities. *See* Sarah Buhr, *Apple Devices Are Butt Dialing 911 from its Refurbishing Facility—20 Times per Day*, TECHCRUNCH (Feb. 22, 2018, 7:54 PM), https://perma.cc/6JJE-2JEL. Notably, there were many incidents in which individuals claim that saying "Siri, call 911" has alerted enforcement agencies, and thereby saved their lives. *See, e.g.*, Yoni Heisler, *College Student Rescued After Using 'Hey Siri" to Call 911 from a Sinking Car*, BGR (Dec. 12, 2019, 3:33 PM), https://perma.cc/FAC8-F3KU.

314. For instance, if an Apple Watch Series 4 or later detects a "hard fall," the watch will help the user in connecting to emergency services. *See Use Fall Detection with Apple Watch*, APPLE, https://perma.cc/32FC-MAMH*; see also*; Elisha Fieldstadt, *Apple Watch's 'Hard Fall' Feature Automatically Calls 911 for Hiker Stranded on Cliff*, NBC NEWS (Oct. 23, 2019, 9:52 AM), https://perma.cc/E43R-KQ9C. When the user does not respond and the watch detects that the user has been immobile for a minute, emergency services will be contacted automatically while sending the location of the user. *Id.*

315. These features are likely to expand to other devices as well, e.g., a smart smoke alarm that contacts the users or emergency services when it registers an unusual or risky heat spike, etc. These solutions, however, are obviously partial, as they rely on user's awareness and knowledge of the existence of such services and how to operate them, but it is mainly limited in the way that users might be unable to even talk during the perilous situation or decide not to report it, and that current technology is limited in the criminal activities it might compute. *See Alexa Can Help Guard Your Home,* AMAZON (2020), https://perma.cc/R8BL-3RNB; *Intruder Alert*, AMAZON (2020), https://perma.cc/6P5T-B7HR ("If you think there is an intruder in your house this skill uses Alexa to make them think twice and encourages them to leave. Alexa pretends to turn on audio and video recording and also pretends to call the Emergency Services.").

316. *See* Russell, *supra* note 196 (discussing Facebook's reaction to the use of its livestreaming service to broadcast a mass shooting).

activity).[317] Recently, Facebook also implemented a "one strike" rule in response to livestreaming crimes, i.e., that those who break its rules will be restricted from using its "Live" feature for a set period of time.[318]

But regardless of how platforms are already responding to practices like livestreaming crime, the role of users in platform governance is a rather limited one.[319] While a court acknowledged the importance of granting immunity for online platforms under § 230,[320] it failed to address the pragmatic barriers for users' participation in the so-called "self-policing" of undesired content.[321] While many scholars argue that the most dominant factor in self-regulation is the economic incentive of platforms to encourage users' engagement in the platform, platforms' decisions will also take other considerations into account.[322]

---

317. Community Standards, *supra* note 104.

318. Russell, *supra* note 196

> Facebook is cracking down on its livestreaming service after it was used to broadcast the shocking mass shootings that left 50 dead at two Christchurch mosques in New Zealand in March. The social network said today that it is implementing a 'one strike' rule that will prevent users who break its rules from using the Facebook Live service.

319. *See* Jack M. Balkin, *2016 Sidley Austin Distinguished Lecture on Big Data Law and Policy: The Three Laws of Robotics in the Age of Big Data*, 78 OHIO ST. L.J. 1217, 1227–31 (2017) (comparing the platform-user relationship to that of a fiduciary and noting that fiduciary relationships involve asymmetries of power, information and transparency); Balkin, *Free Speech in the Algorithmic Society*, *supra* note 194, at 1160

> The fiduciary collects sensitive information about the client that might be used to the client's disadvantage. The client is relatively transparent to the fiduciary, but the fiduciary is not transparent to the client. By this I mean that the client is not well-equipped to understand and monitor the fiduciary's operations. Moreover, the client relies on the fiduciary to perform valuable services, which the client cannot easily perform for themselves.

320. *See* Zeran v. Am. Online, Inc., 129 F.3d 327, 330 (4th Cir. 1997) (barring claims against online service provider under § 230 because defendant was only the publisher, and not the creator, of the tortious content).

321. *See* Klonick, *supra* note 274 at 1608–09.

322. *See* Citron & Norton, *supra* note 303, at 1454 (explaining that platforms are motivated by concerns about the potential business, moral, and

Platforms' potential role as digital Samaritans, whether addressing livestreaming crimes or artificial bystanderism, will be influenced not just by their users' attitudes but also by their own perspective of whether their actions constitute an optimal form of business and management.[323] Putting surveillance capitalism aside, some platforms might self-regulate due to notions of social and corporate responsibility, which are increasingly on the rise.[324] However, eventually, companies and platforms might not respond to users' desires for other reasons, e.g., because the company operates in a monopolistic or an oligopolistic market, users lack the ability to convey their desires effectively to platforms, or users' desires are in conflict with the platforms and their shareholders' own desires.[325] In other words, a variety of market failures could make users' desires less important for some platforms.

While the law will likely not play a direct role in regulating digital Samaritans, it might still play an indirect role in the shaping of platform governance, which will in turn affect the shaping of the algorithmic society.[326] The law still governs how these platforms operate in general, and it also shapes human

---

instrumental costs of digital hate); Klonick, *supra* note 274, at 1615, 1625–30 ("Platforms create rules and systems to curate speech out of a sense of corporate social responsibility, but also, more importantly, because their economic viability depends on meeting users' speech and community norms.").

323. *See* Klonick, *supra* note 274, at 1615 (reasoning that platforms may opt to self-regulate despite § 230 immunity because some see self-regulation as an optimal form of business and company management).

324. *See* Citron & Norton, *supra* note 303, at 1455 ("Some intermediaries are motivated to address digital hate based on their sense of their own corporate social responsibility."); Klonick, *supra* note 274, at 1615 ("[P]latforms have created a voluntary system of self-regulation because they are economically motivated to create a hospitable environment for their users in order to incentivize engagement.").

325. *See* Klonick, *supra* note 274, at 1668 (explaining that users are simply dependent on the whims of corporations, as shareholder values of maximizing company profits are not generally matched with user concerns over equal access and democratic accountability).

326. *See* Balkin, *The Three Laws of Robotics in the Age of Big Data*, *supra* note 319, at 1219 (defining the Algorithmic Society as "a society organized around social and economic decision-making by algorithms, robots, and AI agents").

conduct in various means, both directly and indirectly.[327] One example of indirect regulation in the context of Samaritan duties is that of education.[328] Perhaps as part of a general movement for a better digital education altogether, the state has acknowledged the importance of educating children, students, and military personnel, among others, to aid others in distress under what is termed as "Bystander Intervention Programs."[329] In that way, the state attempts to shape social norms regarding how individuals should act in society, regardless of affirmative duties to report or assist, and respectively, regardless of how platforms govern.[330]

Policymakers might also nudge companies to aid them in law enforcement by creating voluntary mechanisms of reporting by AI operators, as mentioned. They could, for instance, create an incentive for companies to report crimes.[331] Other forms of incentives could also be implemented towards users—encouraging them to act.[332] Others might argue that platforms regulate conduct as a method to preempt regulation, i.e., if these platforms fear that policymakers might impose

327. *See* Balkin, *Free Speech in the Algorithmic Society*, *supra* note 194, at 1179 ("Often it is not even necessary for [governments] to threaten [platforms] directly. Jawboning sends the message that infrastructure providers should be patriotic and cooperate . . . Public officials may also appeal to the public to put pressure on infrastructure providers.").

328. *See* Swan, *supra* note 34, at 981 ("Recently, thousands of schools, college campuses, military bases, workplaces, and other institutions have implemented bystander intervention training programs. These programs are meant to address and prevent social harms like bullying, sexual misconduct, and harassment.").

329. For more on these programs, see *id.* at 981–84.

330. *See id*. at 983 (arguing that bystander intervention programs "try to change social norms so that people are more likely not to look the other way when others are in danger"). Notably, these intervention programs go back to the question of liberty and autonomy within the no-duty-to-rescue common law rule, thus creating "a competing norms problem." *Id*. at 1003. For more drawbacks of bystander intervention programs, see *id*. at 1003–06.

331. Some argue that crafting such incentive program for humans would be morally inappropriate. *See* Kang, *supra* note 60, at 382 (asserting that such incentives for crime deterrence has led to the neglect of "traditional, culpability-based limits in criminal punishment, thereby eroding the value of fairness and proportionality").

332. *See* Perrin-Smith Vance, *supra* note 19, at 147 (suggesting that the state can incentivize, thus encourage, individuals to voluntary report crimes).

affirmative duties on them that do not align with their business models or other perceptions of their platform, they might respond to policymakers' desires and use self-regulation to achieve similar purposes of otherwise obligatory duties.[333] Here, the law regulates the behavior of platforms, without actual legislation or other official regulatory means.[334]

Perhaps the most important role that policymakers should play is that of better equipping users with proper tools to strengthen their position in negotiation with platforms. They must ensure that the public has relevant knowledge and proper access to data regarding platforms' terms of service or community standards, as it will otherwise be difficult for them to be a part of platform reshaping.[335] As such, while transparency is often associated with various state actions, the new governance powers of platforms must push policymakers to extend transparency requirements far beyond any general public companies' requirements often linked with economic reasons.[336] In cases where policymakers realize that platforms are engaged in active enforcement, society members must have proper transparency over the latter's decisions and actions.

Lacking meaningful regulatory intervention, platform governance will likely shape or reshape Samaritan duties, even if those are not imposed by the "traditional" law. These new Samaritan regulations extend far beyond the regulation of a particular conduct and instead represent how the internet greatly changed human conduct and the role that platforms and users play within the formation of this rather new

333. *See* Klonick, *supra* note 274, at 1615 (summarizing and discussing the forces that motivate private actors to self-regulate).

334. *See id*. (explaining that some platforms decide to self-regulate as an attempt to disincentivize governments from regulating them).

335. *See* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1301–13 (2008) (advocating that platforms should voluntarily take up a commitment to a "technological due process," a model that understands the trade-offs of "automation and human discretion," protects individuals' rights to notice and hearings, and gives transparency to rulemaking and adjudication).

336. For more on current financial transparency requirements of public companies, see *Public Companies*, INVESTOR.GOV, https://perma.cc/8HFW-U4W9.

community.[337] Perhaps it is not the role of the law to educate humans in some contexts, leaving it to the discretion of platforms and users to shape the environment in which they digitally live and operate. Some might even take this argument further, arguing that online behavior of reporting might increase reporting of real-world crimes, thus making bad Samaritan laws unnecessary.[338] Under this argument, platform governance could also shape human conduct in physical interactions.

Platform governance, in this respect, is not solely reserved to the role of rule-enforcers, but rather mostly to the role of rule-makers. When crafting and enforcing their terms of use or community standards, these platforms decide what should constitute proper behavior.[339] Take, for example, Facebook's response to livestreaming crimes via its "one strike" rule, referring to such act as an "offense."[340] The use of the terminology "offense" to describe a violation of community standards—or the "law"—implies how platforms view their role of policy making, actively shaping human conduct.[341] This form

337. For a detailed analysis of platform governance, see Klonick, *supra* note 274, at 1630–69.

338. *See* WALLACE, *supra* note 34, at 192 (stating that bystander effect might not be fully translated into the digital era, as there is evidence to support the conclusion that people tend to behave kindly toward strangers on the internet more so than in real-life situations). *But see* Kaufman, *supra* note 18, at 1390 ("Due to the advent of social media and mobile devices, however, fourth parties also should carry an obligation to report, perhaps as much of a responsibility as third parties."); Benzmiller, *supra* note 18, at 960 (explaining that individuals in jurisdictions where duties to rescue are legally required are more likely to regard such duties as morally required).

339. Platforms will likely include their right to enforce the policies that they formed. *See, e.g.*, *Fitbit Privacy Policy*, *supra* note 228 ("[W]e may use the information to . . . protect against fraud and abuse, respond to a legal request or claim, conduct audits, and enforce our terms and policies.").

340. Facebook Vice President of Integrity Guy Rosen was quoted saying "[f]rom now on, anyone who violates our most serious policies will be restricted from using Live for set periods of time—for example 30 days—starting on their *first offense*. For instance, someone who shares a link to a statement from a terrorist group with no context will now be immediately blocked from using Live for a set period of time." *See* Russell, *supra* note 196 (emphasis added).

341. For more on the potential linguistic effects of associating terms from the criminal realm to describe unauthorized use or civil law violations, see HABER, *supra* note 78, at 142–47.

of platform education must not be waived quickly, as it could shape individuals' perceptions of governance and rulemaking.[342]

Platforms' decisions also shape much of our free speech and privacy rights.[343] But leaving it up to platforms to decide on how human rights and liberties are protected might be dangerous, or at least, prove to be a slippery slope.[344] It would be utopian, to some extent, if all users' preferences were fully translated into community standards—agreeable by both the platforms and all users.[345] But there are many barriers for such a utopian reality, and perhaps mainly, the fact that these platforms generate their community standards by themselves with little influence by their users due to technological momentum,[346] will not likely advance platforms to make good decisions with respect to human rights and liberties.

One of the biggest fears regarding platform governance is that of the lack of procedural safeguards and meaningful oversight and transparency.[347] This is part of the legislative responsiveness, or tradeoff, to the averseness and liberty-restrictive elements of criminal proceedings.[348] But if these rather new informal forms of governance impose penalties that society considers as more aversive than criminal law, then at least similar safeguards must be placed within the

342. *See id.*

343. *See* Balkin, *Free Speech in the Algorithmic Society*, *supra* note 194, at 1153 ("[O]ur practical ability to speak is subject to the decisions of private infrastructure owners, who govern the digital spaces in which people communicate with each other.").

344. *See* Gillespie, *supra* note 293, at 212–13 (suggesting that the gift of safe harbor under § 230 should be paired with public obligations, including parameters for how moderation should be conducted).

345. *See id.* at 209 (explaining that, while the logic underlying § 230 persists, the safe harbor afforded to social media platforms becomes increasingly problematic as they evolve).

346. *See* Klonick, *supra* note 274, at 1631–35 (describing the process by which Facebook established its Community Standards and the inherent difficulties in forming an intricate system of rules).

347. *See* Gillespie, *supra* note 293, at 199–200 (detailing the controversies surrounding content moderation on social media platforms and suggesting potential solutions).

348. *See id.* at 212–16 (advocating that adjustments can be made to § 230 to balance some shared public obligations to go with the generous immunity it has offered to platforms).

governance of platforms.[349] In other words, we need to have some form of due process to protect from arbitrary enforcement of community standards, akin to some extent to those granted by the Constitution.[350] We might also need to realize that the use of online penalties, such as banning users, would undermine the values of free speech, perhaps the most dominant constitutional protection in American law, as users who are banned from services will be unable to exercise their legally protected right.[351]

There should be little doubt that technology will continue to assume more active roles in public enforcement.[352] It will become more reliant on crime-detection and crime-prevention AI technologies.[353] Enforcement agencies might use

349. One might argue in this respect, that the use of criminal law to regulate behavior, even if it is considered a misdemeanor, is considered by society as a more intrusive form of punishment, than those that the platform would impose for similar misconduct. At least on these grounds, the highest sanction these platforms could impose is that of banning the user from the services. Still, although evidence for such a claim is not currently present, many individuals might prefer paying a relatively small fine, even if it is imposed under the somewhat averseness of criminal procedures, than permanently losing their Facebook or other social media account, especially if they are dependent on these services for everyday tasks and social participation.

350. *See* Perrin-Smith Vance, *supra* note 19, at 143 (describing procedural due process within the Constitution).

351. In *Packingham v. North Carolina*, the Supreme Court held that it is unconstitutional under the First Amendment for states to bar registered sex offenders from using online social media platforms. 137 S. Ct. 1730, 1737 (2017).

352. *See, e.g.*, *Cooperation or Resistance?: The Role of Tech Companies in Government Surveillance*, 131 HARV. L. REV. 1722, 1722 (2018) ("[T]echnology companies have become major actors in the world of law enforcement and national security.").

353. Notably, many jurisdictions have begun to use AI-based automated risk scoring systems within the criminal law system. *See* Filippo A. Raso et al., *Artificial Intelligence & Human Rights: Opportunities & Risks*, BERKMAN KLEIN CTR. FOR INTERNET & SOC'Y AT HARV. U. 18 (Sept. 25, 2018), https://perma.cc/Y6S4-ATQ4 (PDF); Sungyong Kang, *In Defense of the Global Regulation of a "Duty to Report Crime,"* 57 WASHBURN L.J. 77, 80 (2018) [hereinafter Kang, *In Defense of the Global Regulation of a "Duty to Report Crime"*] ("To detect and deter crime in the global-digital era, states have no choice but to increasingly rely on private sector actors who are often crime victims or facilitators, instead of frontline law enforcement officers.").

sophisticated firearms and accessories that could monitor and record data on their discharge and use, systems that can detect and notify enforcement agencies when gunshots are fired in public or in other places,[354] or other forms of smart wearables.[355] Many jurisdictions have already started to use AI combined with facial recognition abilities to detect ongoing or future crimes.[356] It is also plausible that enforcement agencies will use AI technology not merely in public places but indoors as well, e.g., to detect gunshots fired within a house.[357] If these technologies are available for use by the public, then it will be difficult to restrict law enforcement agencies from using them as well.[358] Thus, with the advancement of AI technologies, there will be little escape from the strengthening of the technology-enforcement paradigm.

---

354. *See, e.g.*, Chip Cutter, *Companies Roll Out Gunshot Detectors at the Office*, WALL ST. J. (Feb. 19, 2019, 6:30 AM), https://perma.cc/Q2DS-K69B (stating that some corporations worried about workplace shootings are installing gunfire-detection systems in U.S. offices and factories).

355. *See* Colin Neagle, *How the Internet of Things is Transforming Law Enforcement*, NETWORKWORLD (Nov. 3, 2014, 6:33 AM), https://perma.cc/JWM3-CZDM (detailing the ways the Internet of Things is starting to make waves in law enforcement, from "connected guns that remember exactly when and how they were fired to wearable smart devices designed for police dogs").

356. One example is Dubai's "Oyoon" project, where CCTV cameras equipped with AI and facial recognition tools are used to monitor criminal behavior in tourism, traffic, and brick and mortar facilities. *See* Ali Al Shouk, *How Dubai's AI Cameras Helped Arrest 319 Suspects Last Year*, UAE (Mar. 18, 2019, 10:52 PM), https://perma.cc/3XMM-UC8J.

357. *See* Stephen Shankland, *How the Internet of Things Knows Where Gunfire Happens*, CNET (July 27, 2014, 4:00 AM), https://perma.cc/US8W-FY6W ("SST is expanding its services so it can detect not just outdoor shootings, but also indoor incidents . . . .").

358. In 2001, the Supreme Court considered whether aiming an infrared thermal imaging device—not generally available for public use—to determine the amount of heat emanating from the suspect's home violates the Fourth Amendment. The Court held that the use of a technology that is not in general use violates the Fourth Amendment when it yields "details of the home that would previously have been unknowable without physical intrusion." *See* Kyllo v. United States, 533 U.S. 27, 29, 40 (2001)*; see also* Orin S. Kerr, *The Fourth Amendment and New Technologies: Constitutional Myths and the Case for Caution*, 102 MICH. L. REV. 801, 802 (2004) (discussing the implications of the Supreme Court's decision in *Kyllo*).

As the complete exclusion of technology from enforcement efforts is both undesirable and nonpragmatic, the role of the platforms that produce and govern these technologies when applied within the criminal realm must at least be governed as well. Essentially, how policymakers react to platform governance will greatly depend on their own interests, as will be shortly demonstrated through recent occurrences. Policymakers might, and to a certain extent, have already intervened to directly regulate platforms when these accumulate too much power and influence over their users.[359] The potential strengthening of the user-platform relationship might also lead the state to act on its own behalf—reducing the powers of platforms to govern. If companies will continue to fight for their users against the "traditional state,"[360] much like we have witnessed in the so-called Apple-FBI case and others,[361] then we might witness such regulatory backlashes, which has recently been in fact on the move.

Such a regulatory backlash is in fact already on the move. Responsive to various moves by media giants, at least some United States policymakers began to actively seek ways that would restrain or reduce the powers of online platforms, and perhaps especially, that of social media companies.[362] One

359. *See* Klonick, *supra* note 274, at 1650–51 (explaining that platform architecture has been informed by and subject to government interference, which can be through the more direct need to comply with local laws and jurisdictions, or by the more subtle influences of government lobbying and requests).

360. *See id.* at 1664 (arguing that platforms "have often pushed back against government requests for takedown").

361. Under what was termed by many as the Apple-FBI Standoff, Apple refused to aid the FBI in access to an encrypted iPhone used by the San Bernardino mass shooter, due to free speech and privacy concerns of its other users. Eventually, the FBI was able to get assistance from a third party. *See In re* An Apple iPhone Seized During Execution of a Search Warrant on a Black Lexus IS300, Cal. L, No. ED 15-0451M, 2016 WL 618401, at *1–2 (C.D. Cal. Feb. 16, 2016); *In re* Apple, Inc., 149 F. Supp. 3d 341, 349 (E.D.N.Y. 2016); Haber, *The Wiretapping of Things*, *supra* note 213, at 772–73.

362. *See* Jeffrey D. Neuburger, *Commerce Dept. Petitions FCC to Issue Rules Clarifying CDA Section 230*, NAT'L L. REV. (July 30, 2020), https://perma.cc/AHM8-HF39 (discussing the petition submitted by the Commerce Department to the Federal Communications Commission (FCC),

recent move was triggered, at least partially, by a feud between the President of the United States and Twitter, and was directed at curbing the protections afforded by § 230.[363] In response to Twitter's actions, Donald Trump signed an Executive Order (EO) aimed to address what he deemed "online bias" in the form of censorship of opinions not held by the platforms themselves.[364] Under justifications of preserving freedom of expression and public discourse, the EO tasks federal agencies with reinterpreting and clarifying § 230, including its requirement for an action "taken in good faith," so as to narrow the protection granted to online platforms by it, in the advancement of the policy described and the spirit expressed by the EO.[365] The EO further seeks to deny online platforms of federal funds paid for marketing and advertising on online platforms, should the latter engage in speech restrictions and be considered "problematic vehicles for government speech due to viewpoint discrimination, deception to consumers, or other bad practices."[366]

---

which requested the FCC initiate a rulemaking to clarify the provisions of § 230 of the Communications Decency Act).

363. In a series of events, some of President Trump's tweets were labeled as "misleading" and applied a fact-check by the platform, while others were labeled as violating platform policy against abusive behavior and the glorifying of violence and their exposure and further dissemination were restricted. *See* Rishi Iyengar & Donie O'Sullivan, *Twitter Temporarily Restricted Trump Campaign's Ability to Tweet Over False COVID-19 Claims*, CNN BUS. (Aug. 6, 2020), https://perma.cc/96XG-RFQ3. In response, Twitter has removed content and restricted the Presidential campaign's account. *See id.*; Makena Kelly, *Twitter Labels Trump Tweets as 'Potentially Misleading' for the First Time*, VERGE (May 26, 2020, 6:04 PM), https://perma.cc/67NA-U6B8; Donie O'Sullivan, *Twitter Puts Warning on Trump Tweet for 'Threat of Harm' Against DC Protesters*, CNN BUS. (June 23, 2020), https://perma.cc/P25K-2WTC; Jon Porter, *Twitter Restricts New Trump Tweet for 'Glorifying Violence'*, VERGE (May 29, 2020), https://perma.cc/SR4P-VU5A.

364. Exec. Order No. 13925, 28 Fed. Reg. 34079 (June 2, 2020).

365. *Id.*

366. *Id.* Accordingly, and despite the EO being challenged in court as unconstitutional, the Department of Justice has moved forward with recommendations to reform § 230, and the Commerce Department has submitted a petition to the FCC to clarify the section's provisions. *See* Neuburger, *supra* note 362 (detailing the increasingly "turbulent" currents around the Communications Decency Act); Makena Kelly, *Donald Trump*

Another move could be located within the rubric of competition and antitrust law. Following an investigation into competition in the digital marketplace, meant to determine whether the market's dominant platforms had amassed their powers through potentially anti-competitive means,[367] a congressional hearing opened with cautioning against the power

Signs Executive Order Targeting Social Media Companies, VERGE (May 28, 2020, 4:32 PM), https://perma.cc/9HRZ-D6K5 (reporting that President Trump signed an executive order to strip liability protection from companies that censure content following a "spat" between the President and Twitter); *CDT Suit Challenges President's Executive Order Targeting First Amendment Protected Speech*, CTR. DEMOCRACY & TECH. (June 2, 2020), https://perma.cc/M3ZW-GFRT (stating that the Center for Democracy & Technology filed a lawsuit against President Trump's Executive Order on Preventing Online Censorship, arguing that the Executive Order violates the First Amendment by curtailing and chilling the constitutionally protected speech of online platforms and individuals). Several more bipartisan reform bills meant to update § 230 have been introduced and further debated, including the PACT Act, which aims to enhance transparency regarding content moderation and to hold companies accountable for content which violates their own policies or is illegal. *See* U.S. DEP'T OF JUST., SECTION 230—NURTURING INNOVATION OR FOSTERING UNACCOUNTABILITY? KEY TAKEAWAYS AND RECOMMENDATIONS (2020), https://perma.cc/E789-W3CY (PDF) (analyzing § 230 and concluding that the "time is ripe" to realign the scope of the statute with the realities of the modern internet); *Schatz, Thune Introduce New Legislation to Update Section 230, Strengthen Rules, Transparency On Online Content Moderation, Hold Internet Companies Accountable for Moderation Practices*, SENATE.GOV (June 24, 2020), https://perma.cc/PCS9-CEG7 (discussing the aims of the Platform Accountability and Consumer Transparency Act, new bipartisan legislation to update § 230 by strengthening transparency in the process online platforms use to moderate content and hold those companies accountable for content that violates their own legal policies or is illegal).

367. *See* Tony Romm, *Amazon, Apple, Facebook and Google Grilled on Capitol Hill Over Their Market Power*, WASH. POST (July 29, 2020, 6:55 PM), https://perma.cc/BDX9-9UQ9 (stating that Congress opened an investigation of Amazon, Apple, Facebook and Google in 2019 aiming to explore "whether the tech industry's most influential quartet of companies had attained their status through potentially anti-competitive means").

wielded by these platforms,[368] and during which, content moderation practices were also inquired upon.[369]

The third potential move against platform governance could stem from national security aspects. A recent example is that of the popular Chinese-owned TikTok platform that has raised concerns with the Trump Administration regarding the Chinese government's influence over content moderation and access to United States' users' data, and was met with an EO that bars companies from engaging in any transactions with TikTok and its parent company, while citing national security reasons for such move.[370]

With much uncertainty regarding the future of these moves, the regulatory pushback has yet to bear any meaningful fruits with regard to reining in platform governance.[371] Until these and similar efforts come to fruition, platforms will continue to govern users' behavior. Such platform governance will greatly shape how Samaritans, or other actions once solely within the prerogative of the state, are regulated. The moral duties of society will thus be greatly shaped by platforms. As long as users will be limited in conveying their moral attitudes to platforms, and regulators abstain from imposing meaningful

368. *See* Adi Robertson, *Everything You Need to Know From the Tech Antitrust Hearing*, VERGE (July 29, 2020, 1:10 PM), https://perma.cc/S47T-Q22U (detailing the congressional antitrust hearing, where the CEOs of four of the world's biggest tech companies endured a five-plus-hour cross-examination filled with combative and accusatory questioning).

369. *See* Neuburger, *supra* note 362 (discussing the need for provider to clarify their content modification practices).

370. *See* Tyler Sonnemaker, *Trump Just Issued 2 Executive Orders Aimed at Chinese-Owned Apps, Barring US Companies from Doing Business with TikTok Parent Company ByteDance and Messaging App WeChat*, BUS. INSIDER (Aug. 6, 2020, 9:49 PM), https://perma.cc/ZPE2-TU3R; Paige Leskin, *Trump's Push to Ban TikTok in the US, Explained in 30 Seconds*, BUS. INSIDER (Aug. 8, 2020, 12:40 PM), https://perma.cc/87B5-HRVG; Nikki Carvjal & Caroline Kelly, *Trump Issues Orders Banning TikTok and WeChat from Operating in 45 Days If They Are Not Sold by Chinese Parent Companies*, CNN POLITICS (Aug. 7, 2020, 7:26 PM), https://perma.cc/Z25K-C2R7.

371. *See* Jeffrey D. Neuburger, *The Communication Decency Act and the DOJ's Proposed Solution: No Easy Answers*, NAT'L L. REV. (June 19, 2020), https://perma.cc/6NXR-UL32 (stating that while critics of the Communications Decency Act (CDA) raise valid concerns, the CDA's protections are extremely important for organizations that operate websites, mobile apps, social media networks, and other online services).

liability on these platforms, the future of not merely Samaritan laws but public enforcement altogether might be shaped almost solely by these platforms. The impact on human rights and liberties in this context might also be enhanced with the potential use of new virtual worlds, which might require reevaluating how rights and liberties are manifested within these new social environments, in addition to new rights and duties that may be created.[372] Thus, it must constantly be on the agenda of policymakers, the media, and users, to join forces against the current and emerging threats of platform governance to individuals rights and liberties, and the reconstruction of society.

V. Conclusion

AI technology will likely reshape many forms of human conduct, including law enforcement. Law enforcement agencies will continue to use technology to detect criminal activities, as long as the legal framework enables such use. Still, it is highly doubtful that current Samaritan duties, applicable in the physical world, will be fully translated into the digital world, imposing affirmative duties on individuals or AI platforms and services to report or aid those in perilous situations. The negative effects of and pragmatic barriers to such duties are simply too high. But the law is merely one form of behavior regulation, thus the lack of a legal requirement does not rule out the role that online platforms and users are already playing in shaping the moral landscape of communication technologies. The idea behind Samaritan statutes is not dead yet, and the ways in which society and online platforms will react to new emerging threats will have a great impact on the social perception of the online community.

But the most prominent Samaritans in the digital era are likely to be the platforms that either enable livestreaming crimes or have the ability to detect criminal activities or other perilous situations in real time. How they will use their great

372. One example is Facebook purchasing Oculus to grant its users new forms of social media within the realm of virtual reality. *See* Casey Newton, *Oculus Will Add New Social Features Powered by Facebook*, VERGE (Sept. 25, 2019, 2:11 PM), https://perma.cc/CNY8-PAC3.

abilities will depend on various factors—driven mostly by market forces and users' attitudes. The problem, however, is that the powers of such platforms, translated into new forms of governance, already shape the ways human rights and liberties are constructed. When platform governance is generally excluded from the legal framework set to protect these human rights and liberties and lacking proper transparency or meaningful oversight on their rulemaking, the foundations of liberal democracies might be at risk. Without any form of meaningful intervention, either legal or social, platforms will govern more and more aspects of our lives, leading us well into living within an algorithmic state. These platforms will likely become the new Samaritans of the algorithmic society.