PREPARED WRITTEN TESTIMONY AND STATEMENT FOR THE RECORD

OF

Danielle Keats Citron,
*Morton & Sophia Macht* Professor of Law, University of Maryland Carey School of Law[*]

HEARING ON

"The National Security Challenge of Artificial Intelligence, Manipulated Media, and 'Deep Fakes'"

BEFORE THE

House Permanent Select Committee on Intelligence

June 13, 2019
Longworth House Office Building
Room 1100
Washington, D.C.

[*] As of July 1, 2019, Citron will join the faculty of Boston University School of Law as a Professor of Law.

## I. INTRODUCTION

Chairman Schiff, Ranking Member, and Committee Members, thank you for inviting me to appear before you to testify about manipulated media generally and "deep fakes" specifically. My name is Danielle Keats Citron. I am the *Morton & Sophia Macht* Professor of Law at the University of Maryland Carey School of Law where I have taught for fifteen years. On July 1, 2019, I am joining the faculty of Boston University School of Law as a Professor of Law. In addition to my home institutions, I am an Affiliate Scholar at Stanford Law School's Center on Internet & Society, Affiliate Fellow at Yale Law School's Information Society Project, and Tech Fellow at NYU Law's Policing Project. I am a member of the American Law Institute where I have been an adviser to the Restatement Third, *Information Privacy Principles Project*. I have written extensively about privacy, free speech, and civil rights, publishing more than 30 articles in major law reviews and scores of opinion pieces for major news outlets.[1] My book HATE CRIMES IN CYBERSPACE (Harvard University Press 2014) tackled the phenomenon of cyber stalking. In academic and popular writing, Professor Robert Chesney and I have explored the looming challenges to democracy, national security, and privacy posed by deep fakes.[2] My testimony grows out of and draws upon that research.

Pictures may be worth a thousand words, but little is as persuasive as audio and video recordings. Audio and video recordings allow us to become firsthand witnesses to events, obviating the need to trust others' accounts. They let us see and hear for ourselves.[3] Even the Supreme Court has endorsed the truth-telling power of audio and video content: If a video shows someone driving recklessly, then the person drove recklessly.[4] Creators of deep fakes count on us to rely on what our eyes and ears are telling us, and therein lies the danger.

---

[1] *See, e.g.*, *Sexual Privacy*, 128 YALE L.J. 1870 (2019); *When Law Frees Us to Speak*, 87 FORDHAM L. REV. 2317 (2019) (with Jonathon Penney); *Why Sexual Privacy Matters for Trust*, 96 WASH. U. L. REV (forthcoming 2019); *Four Principles for Digital Speech*, 95 WASH. U. L. REV. 1353 (2018) (with Neil Richards); *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 NOTRE DAME L. REV. 1035 (2018); *Risk and Anxiety: A Theory of Data Breach Harms*, 96 TEXAS L. REV. (2018) (with Daniel J. Solove); *The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity*, 86 FORDHAM L. REV. 401 (2017) (with Benjamin Wittes); *The Privacy Policymaking of State Attorneys General*, 92 NOTRE DAME L. REV. 747 (2016); *Spying Inc.*, 72 WASH. & LEE L. REV. 1243 (2015); *Criminalizing Revenge Porn*, 49 WAKE FOREST L. REV. 345 (2014) (with Mary Anne Franks); *The Scored Society*, 89 WASH. L. REV. 1 (2014) (with Frank Pasquale); *The Right to Quantitative Privacy*, 98 MINN. L. REV. 62 (2013) (with David Gray); *Intermediaries and Hate Speech: Fostering Digital Citizenship for the Information Age*, 91 B.U. L. REV. 1435 (2011) (with Helen Norton); *Network Accountability for the Domestic Intelligence Apparatus*, 62 HASTINGS L.J. 1441 (2011) (with Frank Pasquale); *Mainstreaming Privacy Torts*, 99 CAL. L. REV. 1805 (2010); *Government Speech 2.0*, 87 DENVER U. L. REV. 899 (2010) (with Helen Norton); *Fulfilling Government 2.0's Promise with Robust Privacy Protections*, 78 GEO. WASH. L. REV. 822 (2010); *Law's Expressive Value in Combating Cyber Gender Harassment*, 108 MICH. L. REV. 373 (2009); *Cyber Civil Rights*, 89 B.U. L. REV. 61 (2009); *Technological Due Process*, 85 WASH. U. L. REV. 1249 (2008); *Reservoirs of Danger: The Evolution of Public and Private Law at the Dawn of the Information Age*, 80 S. CAL. L. REV. 241 (2007).

[2] *Deep Fakes: The Looming Crisis for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. (forthcoming 2019) (with Robert Chesney); Symposium Foreword, 21st Century Truth Decay: Deep Fakes and the Challenge for Free Expression, National Security, and Privacy, MD. L. REV. (forthcoming 2019) (with Robert Chesney); *Deep Fakes and the New Disinformation War*, FOREIGN AFFAIRS, January/February 2019 edition (with Robert Chesney); *Disinformation on Steroids: The Threat of Deep Fakes*, COUNCIL ON FOREIGN RELATIONS ISSUE BRIEF, Oct. 16, 2018 (with Robert Chesney); *Deep Fakes: A Looming Crisis for National Security, Democracy, and Privacy?*, LAWFARE, February 21, 2018 (with Robert Chesney). I rely extensively on my coauthored work in writing this testimony. I am indebted to my coauthor Bobby Chesney for his insights.

[3] *See* Jennifer L. Mnookin, *The Image of Truth: Photographic Evidence and the Power of Analogy*, 10 YALE J. L. HUM. 1, 1-4 (1998).

4. Scott v. Harris, 550 U.S. 372, 380–81 (2007) (ruling that a videotape of a car chase constituted definitive proof of facts so as to preclude the necessity of a trial on the merits).

At a time when even the most basic facts are in dispute, the persuasiveness of audio and video content might offer welcome clarity. Video and audio recordings, however, can do more to mislead and misdirect than to illuminate and inform. Widely-available technologies now permit the creation of fabricated video and audio recordings. Today, with a laptop, internet access, and some technical skills, anyone can create fairly convincing video and audio fakes. Put simply, what our eyes and ears are telling us may not be true.

The technology behind the fakery is poised to take a great leap forward. Deep-fake technologies will enable the creation of highly realistic and difficult to debunk fake audio and video content. Soon, it will be easy to depict someone doing or saying something that person never did or said. Soon, it will be hard to debunk digital impersonations in time to prevent significant damage. Video and audio fakes will be impossible to distinguish from the real thing.

Yes, disinformation has a long and storied history, but it has urgent relevance today. Deep fakes arrive just as the functioning of the marketplace of ideas is under serious strain. Whereas media outlets committed to professional standards once supplied much of our information diet, social media platforms are increasingly the go-to source for information. Companies tailor content to our interests and views. The resulting digital echo chambers make it seem as if everyone shares our way of thinking.

Falsehoods spread like wildfire on social networks. Social media platforms are susceptible to information cascades, whereby people pass along information shared by others without checking its validity. Information that goes viral tends to be controversial and salacious. People are attracted to, and more likely to spread, negative and novel information.[5] Bots escalate the spread of misinformation.[6]

The declining influence of traditional media, cognitive biases, and social-media information dynamics have already fueled the spread of fake video and audio recordings.[7] Although today's fake video and audio content can be debunked fairly quickly (particularly if creators manipulate recordings of events for which authentic copies exist), real recordings often fail to catch up to salacious fakes. Far too often, the fakery is believed. This will grow worse as more sophisticated deep-fake technology emerges.

The circulation of deep fakes has potentially explosive implications for individuals and society. Under assault will be reputations, political discourse, elections, journalism, national security, and truth as the foundation of democracy. My testimony will outline the risks as well as potential legal and market solutions and their limits.

---

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

6. Robinson Meyer, *The Grim Conclusions of the Largest Ever Study of Fake News*, THE ATLANTIC (Mar. 8, 2018), https://www.theatlantic.com/technology/archive/2018/03/largest-study-ever-fake-news-mit-twitter/555104/ (quoting political scientist Dave Karpf).

[7] Robert Chesney, Danielle Citron, and Quinta Jurecic, *That Pelosi Video: What To Do About 'Cheapfakes' in 2020*, Lawfare (May 29, 2019), https://www.lawfareblog.com/about-pelosi-video-what-do-about-cheapfakes-2020.

## II. DESCRIBING THE THREAT LANDSCAPE

### *A. Harm to Individuals and Companies*

Deep-fake videos can target individuals, damaging their reputations and causing severe emotional distress.[8] Let's begin as the deep-fake trend did—with fabricated sex videos. In late 2017, word got out about a subreddit devoted to deep-fake sex videos of celebrities.[9] At the time, the subreddit (now closed) had countless threads. Redditors posted deep-fake sex videos of Gal Gadot and Emma Watson. They sought help in making deep fakes. One person said he wanted to make a deep-fake sex video of his ex-girlfriend and wondered if 30 photographs would be sufficient. Posters directed one another to YouTube tutorials providing instructions on the creation of deep-fake videos.[10]

Consider the experience of Noelle Martine. Ms. Martine was a high-school student in Australia when she posted videos and photos of herself on social media.[11] Someone used the photos to insert her face into pornographic images; the doctored photos were posted online alongside her home address and cell phone number. Not long thereafter, a deep-fake sex video of Ms. Martine appeared, showing her performing oral sex on a strange man. She was inundated with death and rape threats, and strangers contacted her for sex. Keep in mind that Ms. Martine was a high-school student. Law enforcement told her that nothing could be done about the postings.[12]

Rana Ayyub is a journalist whose reporting has exposed corruption in Hindu national politics.[13] In an apparent effort to silence her, pseudonymous posters circulated a deep-fake sex video featuring her. The video went viral.[14] It was shared via Twitter and text messages. It appeared in posts alongside her home address, phone number and the phrase: "I am available."[15] Her Twitter feed was overwhelmed with screenshots of the video. Death and rape threats filled her email inbox. Ms. Ayyub was terrified. For weeks, she could not write, let alone speak. She could barely eat.[16] She fears having people take photos of her in public lest they use them to create more deep fakes.[17]

---

8. *See* Robert Chesney & Danielle Keats Citron, *Deep Fakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics*, Foreign Aff., Jan.–Feb. 2019, at 153 https://www.foreignaffairs.com/articles/world/2018-12-11/deepfakes-and-new-disinformation-war.

9. Samantha Cole, *AI-Assisted Fake Porn Is Here and We're All Fucked*, MOTHERBOARD (Dec. 11, 2017, 7:18 PM), https://motherboard.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn. For a description of subreddits, see *What are Communities or "Subreddits"?*, REDDIT, https://www.reddithelp.com/en/categories/reddit-101/communities/what-are-communities-or-subreddits (last visited Apr. 6, 2019) ("Reddit is a large community made up of thousands of smaller communities . . . known as 'subreddits.'").

10. *See, e.g.*, tech 4tress, *Deepfakes Guide: Fake App 2 2 Tutorial. Installation (Totally Simplified, Model Folder Included)*, YOUTUBE (Feb. 21, 2018), https://www.youtube.com/watch?v=Lsv38PkLsGU; The Great Zasta, *How to Merge Faces with Fake App in 5 Minutes!! Quickest Tutorial*, YOUTUBE (Feb. 18, 2018), https://www.youtube.com/watch?v=i4bar4X7ghs .

11. *See* Ally Foster, *Teen's Google Search Reveals Sickening Online Secret About Herself*, NEWS.COM.AU, (June 30, 2018), https://www.news.com.au/technology/online/security/teens-google-search-reveals-sickening-online-secret-about-herself/news-story/ee9d26010989c4b9a5c6333013ebbef2.

12. TEDx Talks, *Sexual Predators Edited My Photos into Porn—How I Fought Back, Noelle Martin, TEXxPerth*, YOUTUBE (Mar. 6, 2018), https://www.youtube.com/watch?v=PctUS31px40.

13 *See, e.g.*, RANA AYYUB, GUJARAT FILES: ANATOMY OF A COVER UP (2017).

14 Siobhan O'Grady, *An Indian Journalist Has Been Trolled For Years, Now U.N. Experts Say Her Life Could Be At Risk*, WASH. POST (May 26, 2018).

15. Rana Ayyub, Opinion, *In India, Journalists Face Slut-Shaming and Rape Threats*, N.Y. TIMES (May 22, 2018), https://www.nytimes.com/2018/05/22/opinion/india-journalists-slut-shaming-rape.html.

16. *Id.*

17 Skype Interview of Rana Ayuub (dated May 9, 2019) (notes on file with author).

Like other sexual-privacy invasions, deep-fake sex videos are likely to have a disproportionate impact on women and marginalized communities.[18] As Ms. Martin and Ms. Ayyub explained, being turned into a sex object without consent is terrifying, embarrassing, and life altering. Deep-fake sex videos reduce people to genitalia, breasts, buttocks, and anus, affixing them with a sexual identity not of their making.[19] People have difficulty finding or keeping jobs when deep-fake sex videos appear in searches of their names. They go offline, even though it may hurt their careers.

Deep-fake videos could be used to sabotage corporate CEOs and their companies. Imagine the night before a company's Initial Public Offering, a deep-fake video appears showing the CEO committing a crime. If the deep-fake video is shared widely, the company's stock price may falter and a tremendous amount of money may be lost. Of course, the video could be debunked in a few days, but by that time the damage has already been done.

### *B. Disrupting Democracy: Social Cohesion, Public Safety, and Elections at Risk*

Deep fakes can undermine social cohesion essential for democratic discourse. We have seen the damage caused by low tech impersonations. In 2016, Russia's state-sponsored disinformation operations succeeded in deepening existing social and ideological fissures in the United States. In the name of Black Lives Matters activists, Russian social media accounts shared inflammatory content in an effort to stoke racial tensions.

In the future, the disinformation could come in the form of a deep-fake video of a white police officer shouting racial slurs. A deep-fake video could feature a well-known imam in New York City celebrating an ISIS attack on American soldiers in Afghanistan. Deep fakes could so exacerbate societal divisions that violence ensues.

A century ago, Justice Oliver Wendell Holmes warned of the danger of falsely shouting fire in a crowded theater.[20] The panic that Holmes imagined might be modest compared to the fallout from a perfectly-timed deep fake. A deep-fake video, if inflammatory enough and timed just right, could fuel unrest and violence. Consider Baltimore City residents' grief and anger after the senseless killing of resident Freddie Gray in police custody. Imagine if the day after Mr. Gray's death, a deep-fake video appeared featuring the police chief endorsing the mistreatment of Mr. Gray. If the video was circulated to protestors, there might have been civil unrest resulting in physical violence.

Deep-fake videos and audios could undermine the democratic process by tipping an election. Imagine that the night before the 2020 election a deep fake showed a candidate in a tight race doing something shocking he never did. The deep fake, if spread widely, could alter the election's outcome. The deep fake creator could be a hostile state actor or non-state actors motivated to sway the election a particular way. No matter, the damage would be irreparable. Elections cannot be undone.

---

18. *See* Mary Anne Franks & Ari Ezra Waldman, *Sex, Lies, and Videotape: Deep Fakes and Free Speech Delusions*, 78 MD. L. REV. (2019).

19. *See* Danielle Keats Citron, *Sexual Privacy*, 128 YALE L. J. 1870 (2019).

[20] Schenck v. United States, 249 U.S. 47, 52 (1919) (Holmes, J.).

### *C. Decaying Trust Essential for Democratic Institutions*

Deep-fake videos will risk deepening people's distrust in civic and political institutions. Deep fakes will find an audience primed to believe the worst about lawmakers, police officers, and reporters. Deep-fake videos could feature public officials taking bribes. They could purport to show FBI special agents discussing ways to abuse their authority. They could feature journalists saying they made up a politically-divisive story. All of this could spark outrage at the institutions that serve as the foundation of our democracy.

Deep fakes may undermine journalism in other ways. News organizations may be chilled from rapidly reporting real, disturbing events for fear that the evidence will turn out to be fake. One can expect that people will try to trap news organizations in this way. We have already seen stings pursued without the benefit of deep-fake technology. Convincing deep fakes may make such stings more likely to succeed, leaving news organizations fearful to publish video and audio content in the future. Without a quick and reliable way to authenticate video and audio, the press may find it difficult to fulfill its ethical and moral obligations to the truth.

### *D. Endangering National Security and Diplomacy*

Hostile state and nonstate actors will surely leverage deep fakes to accomplish their goals. Terrorist organizations, for instance, could distribute deep fakes depicting adversaries—government officials, military officers, or soldiers—engaging in provocative actions, with content chosen to maximize the galvanizing impact on target audiences. Deep fakes could purport to show an Army general burning a copy of the Koran or U.S. soldiers murdering children. Such fakes could lead to violent reprisals and aid terrorist recruitment.

Diplomacy is at risk as well. What if in the middle of sensitive negotiations, diplomats are shown a deep fake suggesting an adversary is disingenuous? The video could scuttle diplomatic efforts. Deep fakes could be used in counterintelligence operations, inspiring reprisal against U.S. assets and agents.

### *E. Escaping Accountability for the Truth: The Liar's Dividend*

It is not just that deep fake content can wreak havoc by spreading lies. Wrongdoers could invoke the possibility of deep fakes to escape the truth.

In an environment of pervasive, highly-realistic deep fakes, people caught in genuine recordings of misbehavior will find it easier to cast doubt on damning evidence. They could dismiss *real* video and audio as "deep fakes." Education about deep fakes could inure to the benefit of wrongdoers in giving the public a sound reason to disregard real videos or audios.[21] Professor Robert Chesney and I have called this the "Liar's Dividend."

President Trump has tried to leverage this possibility in denouncing the veracity of the Lester Holt interview (where he admitted that he fired FBI Director James Comey because of that "Russia

[21] See Chesney & Citron, *supra* note.

issue") and the Access Hollywood tape (where he said when you are a star you can "grab 'em by the pussy").[22] If the public were more sensitized to the deep-fake phenomenon, his assertions might have been believed.

The more people are educated about the advent of deep fakes, the more they may disbelieve real recordings. Regrettably and perversely, the Liar's Dividend grows in strength as people learn more about the dangers of deep fakes. But that is not to give up on the project of educating the public about deep fakes, but rather to note that education efforts must include warnings about the Liar's Dividend.

## III. SKETCHING LEGAL SOLUTIONS AND CHALLENGES

Unfortunately, there are no easy answers to these concerns. Law could mitigate some, but only some, of the threat. A legal agenda also faces significant practical hurdles. We need the law, tech companies, and a heavy dose of societal resilience to make our way through these challenges.

### *A. Legal Tools and Free Speech Considerations*

No criminal or civil liability regime specifically addresses the creation or distribution of deep fakes. A ban on deep fake technology would not be desirable. Digital manipulation is not inherently problematic. There are pro-social uses of the technology. Deep fakes exact significant harm in certain contexts but not in all.[23]

Existing civil and criminal laws would address certain deep fakes. Tort law would provide redress for some deep-fake scenarios. Deep-fake creators could be sued for defamation where falsehoods are circulated recklessly in the case of public figures or officials or negligently in the case of private individuals. The "false light" tort—recklessly creating a harmful and false implication about someone in a public setting—likewise has potential for certain cases. Subjects of deep fakes may be able to bring claims for intentional infliction of emotional distress, which requires proof of "extreme and outrageous conduct." Public figures could bring "right of publicity" claims if defendants generate financial gain from the fakes.[24]

Criminal law offers limited avenues for deterrence and punishment. A handful of states criminalize impersonations that cause certain injuries. In a few jurisdictions, creators of deep fakes could face charges for criminal defamation if they posted videos knowing they were fake or if they were reckless as to their truth or falsity.[25] If perpetrators post deep fakes in connection

[22] In public congressional testimony in the winter of 2019, DNI Director Dan Coats and CIA Director Gina Haspel expressed their disagreement with the President's policy towards Syria and ISIS. President Trump responded swiftly to the video. Rather than criticizing them, the President simply asserted that neither Coats nor Haspel had disagreed with him. President Trump said, in so many words, that it was all fake news—that the officials said that they agreed with his policies.

[23] It may be possible to draft a federal criminal law banning deep fakes in a sufficiently narrow way that would withstand judicial scrutiny. Senator Ben Sasse has proposed a federal criminal statute that would extend to creators and publishers of deep fakes (including online platforms) if they knew that the deep-fake content is fake and published a deep fake knowing it enabled crimes or torts.

[24] *See generally* JENNIFER ROTHMAN, RIGHT OF PUBLICITY: PRIVACY REIMAGINED FOR A PUBLIC WORLD (2018).

25. *See* Eugene Volokh, *One to One Speech Versus One-to-Many Speech*, 107 NW. U. L. REV. 731 (2013).

with the persistent targeting of individuals, they might be prosecuted for violating the federal cyberstalking law as well as analogous state statutes.

Deep fakes implicate freedom of expression, even though they involve intentionally false statements. In *United States v. Alvarez*,[26] decided in 2012, plurality and concurring opinions of the Supreme Court concluded that "falsity alone" does not remove expression from First Amendment protection.[27] As the plurality noted, falsehoods generally warrant protection because they inspire rebuttal and "reawaken respect" for valuable ideas in public discourse.[28] Central to this point is faith in the public's willingness to counter lies and engage in reasoned discourse. All of the Justices, however, agreed that false factual statements could be regulated in the presence of harm, but differed in the particulars.[29]

The legal approach outlined here would comport with First Amendment commitments. To start, certain categories of speech are not covered by the First Amendment due to their propensity to bring about serious harms and their slight contribution to free speech values.[30] Some deep fakes will fall into those categories and thus provide the basis for legal restrictions. This includes defamation of private persons, fraud, true threats, and the imminent-and-likely incitement of violence.[31] Speech integral to criminal conduct like extortion, blackmail, and perjury has long been understood to enjoy no First Amendment protection.[32]

The First Amendment would likely countenance prosecutions for harm-causing impersonations of individuals. As Helen Norton explains, laws banning the impersonation of government officials are "largely uncontroversial as a First Amendment matter in great part because they address real (if often intangible) harm to the public as well as to the individual target."[33]

Free expression values would not be undermined. Lies about the source of speech—whether a particular person is the one actually speaking—often undermine, rather than protect, free speech values.[34] Deep fakes deny listeners the ability to assess the quality and credibility of speech, undermining democratic self-governance and the search for truth.[35] They undermine trust as to who is actually speaking and make it difficult to assess speakers' reliability. To be sure, law should not proscribe deep fakes that amount to parody and satire, and the legal agenda advocated here would not do so.

---

26. United States v. Alvarez, 567 U.S. 709 (2012) (plurality opinion).
27. *Id.* at 719.
[28] *Id.* at 719, 722.
[29] *Alvarez*, 567 U.S. at 719 (plurality opinion), 731-34 (Breyer, J., concurring), and 750 (Alito, J. dissenting).
30. *See generally* CITRON, HATE CRIMES IN CYBERSPACE, *supra* note **Error! Bookmark not defined.**, at 199-218 (discussing narrow categories of low-value speech accorded less rigorous protection or no protection under First Amendment analysis).
31. *See* Alan Chen & Justin Marceau, *High Value Lies, Ugly Truths, and the First Amendment*, 68 VAND. L. REV. 1435 (2015).
32. CITRON, HATE CRIMES IN CYBERSPACE, supra note, at 203-05.
[33] Helen Norton, *Lies to Manipulate, Misappropriate, and Acquire Government Power*, in LAW AND LIES 143, 168 (Austin Sarat ed., 2015).
[34] *Id.* at 168.
[35] Helen Norton, *Thirteen Ways of Looking at Election Lies*, 71 OKLA. L. REV. 117, 131 (2018).

*B. Limits*

In some contexts, legal claims and criminal prosecutions may be theoretically possible but practically infeasible. For a start, it may be difficult to attribute the creation of a deep fake to a particular person or group. For civil claims and criminal prosecutions to work, perpetrators need to be found and identified.

Even if perpetrators are identified, they may be beyond the court's reach, as in the case of foreign individuals or governments. To pursue actions against creators, courts need to have jurisdiction over them. Then too, criminal investigations may founder if officers lack training in technology and the law.[36]

Civil litigation may not be a practical response given its expense. Individuals usually bear the costs of bringing civil claims, and those costs can be steep. For most people, it would be too costly to sue deep-fake creators. It would be difficult to find lawyers to work on cases on a contingency basis because many creators will be judgment-proof.

To be sure, lawyers would have an incentive to take cases on contingency if social media providers could be sued. Although online platforms are in the best position to minimize harm, they enjoy a broad-sweeping immunity from liability for user-generated content under Section 230 of the Communications Decency Act. The immunity means that platforms are free to ignore the propagation of damaging deep fakes, even ones that platforms know cause specific and immediate harms.

As I have argued with Benjamin Wittes,[37] Quinta Jurecic,[38] and Professor Chesney,[39] that federal immunity should be amended to condition the immunity on reasonable moderation practices rather than the free pass that exists today. The current interpretation of Section 230 leaves platforms with no incentive to address destructive deep-fake content. To be sure, there are platforms that do not need civil liability exposure to combat such obvious harms; market pressures and morals in some cases are enough. However, market pressures and morals are not always enough, and they should not have to be.

I am grateful to Chairman Schiff for inviting me to discuss my research with Professor Robert Chesney. I appreciate the Committee's engagement on this issue and its desire to tackle the looming challenges raised by deep fakes.

[36] As I explore in my book, law enforcement routinely fails to address cyber stalking and other forms of online abuse for these reasons. Citron, supra note.
[37] *See* Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity*, 86 FORDHAM L. REV. 401, 407 n.52 (2017).
[38] Danielle Citron & Quinta Jurecic, *Platform Justice: Content Moderation at an Inflection Point*, HOOVER INST. (2018), https://www.hoover.org/sites/default/files/research/docs/citron-jurecic_webreadypdf.pdf [https://perma.cc/V7T4-X8Y4].
[39] Chesney & Citron, *supra* note.

**BIOGRAPHY**

On July 1, 2019, Professor Danielle Citron will join the faculty of Boston University School of Law as a Professor of Law. From 2004 to June 30, 2019, she taught at the University of Maryland Carey School of Law where she received the 2018 "UMD Champion of Excellence" award for teaching and scholarship. Professor Citron has been a Visiting Professor at Fordham University School of Law (Fall 2018) and George Washington Law School (Spring 2017). After settling in at BU Law, she will visit Harvard Law School. Professor Citron teaches and writes about data privacy, free expression, civil rights, and administrative law.

Professor Citron is an internationally recognized privacy expert. Her book *Hate Crimes in Cyberspace* (Harvard University Press) explored the phenomenon of cyber stalking and the role of law and private companies in combating it. The editors of *Cosmopolitan* included her book in its "20 Best Moments for Women in 2014." Professor Citron has published numerous book chapters and more than 30 law review articles, published in the Yale Law Journal, California Law Review (twice), Michigan Law Review (twice), Harvard Law Review Forum, Boston University Law Review (three times), Notre Dame Law Review (twice), Fordham Law Review (twice), George Washington Law Review, Minnesota Law Review, Texas Law Review, Washington University Law Review (three times), Southern California Law Review, Washington & Lee Law Review, Wake Forest Law Review, Washington Law Review (twice), UC Davis Law Review and other journals. Her current scholarly projects concern sexual privacy; privacy and national security challenges of deep fakes; and the automated administrative state.

Professor Citron's opinion pieces have appeared in major media outlets, including *The New York Times, The Atlantic, Slate, Time, CNN, The Guardia*n, *New Scientist, Lawfare, ars technica,* and *New York Daily News*. She is a technology contributor for Forbes and served as a long-time member of the now-defunct Concurring Opinions blog (2008-2019).

Professor Citron's work has been recognized at home and abroad. In 2015, the United Kingdom's Prospect Magazine named Professor Citron one of the "Top 50 World Thinkers." The Maryland Daily Record named her one of the "Top 50 Most Influential Marylanders." In 2011, Professor Citron testified about misogynistic cyber hate speech before the Inter-Parliamentary Committee on Anti-Semitism at the House of Commons.

Professor Citron is an active member of the cyber law community. She is an Affiliate Scholar at the Stanford Center on Internet and Society, Affiliate Fellow at the Yale Information Society Project, Senior Fellow at Future of Privacy, and Tech Fellow at the NYU Policing Project. She is a member of the American Law Institute (inducted in 2017) and serves as an adviser to the American Law Institute's Restatement Third, *Information Privacy Principles Project*. She is a member of the Principals Group for the Harvard-MIT AI Fund. Professor Citron works with civil liberties and privacy organizations. She is the Vice President of the *Cyber Civil Rights Initiative*. She served as the Chair of the *Electronic Privacy Information Center*'s Board of Directors from 2017-2019 and now sits on its Board. Professor Citron has served on the Advisory Boards of *Without My Consent*, *Teach Privacy*, *SurvJustice*, and the *International Association of Privacy Professionals* Privacy Bar. In connection with her advocacy work, she advises tech companies

on online safety, privacy, and free speech. She serves on Twitter's Trust and Safety Council as well as Facebook's Nonconsensual Intimate Imagery Task Force. She has presented her research at Twitter, Facebook, Google, and Microsoft.

Professor Citron advises federal and state legislators, law enforcement, and international lawmakers on privacy issues. In July 2017, she testified at a congressional briefing on online harassment and sexual violence co-sponsored by Congresswoman Jackie Speier. In April 2015, she testified at a congressional briefing sponsored by Congresswoman Katharine Clark on the First Amendment implications of a federal cyber stalking legal agenda. She has worked with the offices of Congresswoman Katharine Clark, Senator Elizabeth Warren, Senator Kamala Harris, and Senator Diane Feinstein on federal legislation. Professor Citron helped Maryland State Senator Jon Cardin draft a bill criminalizing the nonconsensual publication of nude images, which was passed into law in 2014. From 2014 to December 2016, Professor Citron served as an advisor to California Attorney General Kamala Harris. She served as a member of AG Harris's Task Force to Combat Cyber Exploitation and Violence Against Women. In October 2015, Professor Citron, with AG Harris, spoke at a press conference to discuss the AG office's new online hub of resources for law enforcement, technology companies, and victims of cyber sexual exploitation.

Professor Citron has presented her research in over 200 talks at federal agencies, meetings of the National Association of Attorneys General, the National Holocaust Museum, the Anti-Defamation League, Wikimedia Foundation, universities, companies, and think tanks. She appeared in HBO's *Swiped: Hooking Up in the Digital Age* (directed by Nancy Jo Sales) and *Netizens* (which premiered at the 2018 Tribeca Film Festival, directed by Cynthia Lowen). She has been quoted in hundreds of news stories in publications including *The New York Times*, *Washington Post*, *Wall Street Journal*, *Los Angeles Times*, *San Francisco Chronicle*, *USA Today*, *National Public Radio*, *Time*, *Newsweek*, *the New Yorker*, *New York Magazine*, *Cosmopolitan*, HBO's John Oliver Show, *Barron's*, *Financial Times*, *The Guardian*, *Vice News*, and *BBC*. She is a frequent guest on National Public Radio shows, including All Things Considered, WHYY's Radio Times, WNYC's Public Radio International, Minnesota Public Radio, WYPR's Midday with Dan Rodricks, Wisconsin Public Radio, WAMU's 1A, WAMU's The Diane Rehm Show, and Chicago Public Radio. She will be giving a TED talk on the issue of deep fakes at this year's Global TED Summit in Edinburgh, Scotland.