PREVENTING INTIMATE IMAGE ABUSE VIA PRIVACY-PRESERVING ANONYMOUS CREDENTIALS

*Janet Zhang and Steven M. Bellovin**

ABSTRACT

The problem of non-consensual pornography (NCP), sometimes known as intimate image abuse or revenge porn, is well known. Despite its distribution being illegal in most states, it remains a serious problem, if only because it is often difficult to prove who uploaded the pictures. Furthermore, the Federal statute commonly known as Section 230 generally protects Internet sites such as PornHub from liability for content created by their users; only the users are liable, not the sites.

One obvious countermeasure would be to require Internet sites to strongly authenticate their users, but this is not an easy problem to solve. Furthermore, while strong authentication would provide accountability for the immediate upload, such a policy would threaten the ability to speak anonymously, a vital constitutional right. Also, it often would not help identify the original offender—many people download images from one site and upload them to another, which adds another layer of complexity.

We instead propose a more complex scheme, based on a privacy-preserving cryptographic credential scheme originally devised by researcher Jan Camenisch and Professor Anna Lysyanskaya. While the details (and the underlying mathematics!) are daunting, the essential properties of their scheme are straightforward. Users first obtain a primary credential from a trusted identity provider; this provider verifies the person's identity, generally via the usual types of government-issued ID documents, and hence knows a user's real identity. To protect privacy, this primary credential can be used to generate arbitrarily many anonymous but provably valid sub-credentials, perhaps one per web site; these sub-credentials cannot be linked to each other or to the primary credential. For technical reasons, sub-credentials cannot be used directly to digitally sign images. Instead, they are used to obtain industry-standard cryptographic "certificates" that can be so used. The certificate-issuing authority also

---

* Janet Zhang holds a B.S. degree from the Columbia University Fu Foundation School of Engineering and Applied Science. Steven M. Bellovin is the Percy K. and Vida L.W. Hudson Professor of Computer Science at Columbia University and affiliate faculty at Columbia Law School. We are grateful for the detailed comments supplied by Zartosht Ahlers, Katerina Kaganovich, Sunoo Park, Daniel Richman, and Rebecca Wexler.

receives and retains an encrypted, random pseudonym known by the identity provider, which is used to identify the web site user. If NCP is alleged to be present in an image, information extracted from the image's metadata plus the encrypted pseudonym can be sent to a deanonymization agent, the only party who can read it. The final step to reveal the uploader's identity is to send the decrypted pseudonym to the identity provider; which knows the linkage between the pseudonym and a real person. In other words, three separate parties must cooperate to identify someone.

The scheme is thus privacy-preserving, accountable, and abuse-resistant. It is privacy-preserving because sub-credentials are anonymous and not linkable to anything. It provides accountability, because all images are signed before upload and the identity of the original uploader can be determined if necessary. It is abuse-resistant, because it requires the cooperation of those three parties—the certificate issuer, the deanonymization agent, and the identity provider—to identify an image uploader. The paper contains a reasonably detailed description of how the scheme works technically, albeit without the mathematics.

Our paper describes the necessary legal framework for this scheme. We start with a First Amendment analysis, to show that this potential violation of the constitutional right to anonymity is acceptable. We conclude that exacting scrutiny (as opposed to the generally higher standard of strict scrutiny), which balances different rights, is the proper standard to use; it is what the Supreme Court has used in, e.g., Citizens United*, to justify violations of anonymity. Here, the balance is the right to anonymous publication of images versus the right to intimate privacy, a concept that we show has also been endorsed by the Supreme Court. We go on to discuss the requirements for the different parties—e.g., their trustworthiness and if they are in a jurisdiction where aggrieved parties would have effective recourse—and the legal and procedural requirements, including standing, for opposing deanonymization. We suggest that all three parties should have the right to challenge deanonymization requests, to ensure that they are valid. We also discuss how to change Section 230 in a way that would be constitutional (it is unclear if use of this scheme can be mandated) but would still induce Internet sites to adopt it. Finally, we discuss other barriers to adoption of this scheme and how to work around them: not everyone will have a suitable government-issued ID, and some sites, especially news and whistleblower sites, may wish to eschew strongly authenticated images to protect the identities of their sources.*

Table of Contents

I. INTRODUCTION

The Internet has had a profound effect on society. One of the most notable areas has been free speech: it provides a platform to those who did not have one. Indeed, a Federal court noted that "It is no exaggeration to conclude that the Internet has achieved, and continues to achieve, the most participatory marketplace of mass speech that this country—and indeed the world—has yet seen."[1] That said, this free-flowing communication is not without its disadvantages, as demonstrated by the concerns that led to the *Communications Decency Act of 1996*,[2] which was passed because Congress was concerned about inappropriate material on the Internet being too readily available to minors.[3]

One of the more troubling forms of content on the Internet is known as "nonconsensual pornography," defined as "the distribution of sexually graphic images of individuals without their consent,"[4] and sometimes, albeit incorrectly, termed "revenge porn."[5] Typically, a male posts intimate pictures of a current or former female sexual partner, though some cases involve same-sex relationships;[6] often, these pictures had been taken consensually as part of an ongoing relationship,[7] but hidden cameras are far from unknown.[8]

---

[1] *ACLU v. Reno*, 929 F. Supp. 824, 881 (E.D. Pa. 1996).

[2] Communications Decency Act, 47 U.S.C. § 223 (1996). These provisions were overturned by the Supreme Court the following year, which held that they violated the First Amendment (*Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997).).

[3] The law criminalized, among other things, "any comment, request, suggestion, proposal, image or other communication which is obscene or indecent, knowing that the recipient of the communication is under 18 years of age". 47 U.S.C. § 223(a)(1)(B).

[4] *See* https://www.cybercivilrights.org.

[5] *See* https://cybercivilrights.org/faqs/. A study has shown that most perpetrators claim that they were not motivated by revenge or anger (Asia A. Eaton et al., *2017 Nationwide Online Study of Nonconsensual Porn Victimization and Perpetration* 19 (Jun. 2017).

[6] A study has shown that women are 1.7 times as likely to have been victimized by or NCP as are men (*Id.* at 12.). Men are more than twice as likely to perpetrate NCP (*Id.* at 15).

[7] Amanda Lenhart et al., *Nonconsensual Image Sharing: One in 25 Americans Has Been a Victim of "Revenge Porn"* 5 (Data & Society Research Institute Dec. 2016), https://datasociety.net/pubs/oh/Nonconsensual_Image_Sharing_2016.pdf (a far higher rate of self-identified LGB Internet users have been victimized than heterosexuals). Note that the rate of victimization reported in Eaton et al., *supra* note 5, is far higher than in Lenhart.

[8] One of the better-known cases is the tragedy of Tyler Clementi, who committed suicide after his roommate used a webcam to secretly watch him having sex with another man; *see* Parker, Ian, *The Story of Suicide*, NEW YORKER (Jan. 30, 2012), https://www.newyorker.com/magazine/2012/02/06/the-story-of-a-suicide.

The question is what to do about the problem. Almost all jurisdictions already criminalize the practice.[9] Establishing accountability, however, is another issue. Under certain circumstances, it may be possible to learn who uploaded a given image to a web site; however, once such images are first uploaded, they may be copied to other sites by third parties who may be completely unaware that the image in question was uploaded without the consent of the subject.[10]

The problem is exacerbated by a statute, originally enacted as part of the Communications Decency Act[11] but today commonly known simply as Section 230, that broadly immunizes web sites for carriage of user-contributed content.[12] Sites are not liable if they host nonconsensual pornography; they are certainly not required to ascertain images ' provenance.[13]

Professor Danielle Citron has suggested that web sites wishing the full protection of Section 230 should log information to identify who uploaded imagery.[14] As we show in Part V.A, *infra*, simple mechanisms for such logging are not likely to suffice. We need to go further.

One obvious solution is to require that all images be tagged with information indicating who created and/or uploaded them. That way, if an image were later deemed to be nonconsensual pornography, the offending party could easily be identified and prosecuted. Such a requirement could run afoul of the First Amendment.[15] The Supreme Court has held in no uncertain terms that there is a right to anonymous publication: "There can be no doubt that such an identification requirement would tend to restrict freedom to distribute information and thereby freedom of expression. 'Liberty of circulating is as essential to that freedom as liberty of

---

[9] According to the Cyber Civil Rights Initiative, 48 states, the District of Columbia, Guam, and Puerto Rico have statutes in place; *see* https://www.cybercivilrights.org/revenge-porn-laws/.

[10] *See, e.g.,* Danielle Keats Citron and Mary Anne Franks, *Criminalizing Revenge Porn,* 49 Wake Forest L. Rev. 345, 360 (2014). ("Once an image is released, getting it removed from one site does not mean that it will be removed from every other site to which it has migrated.")

[11] *See* note 2, *supra*.

[12] 47 U.S.C. § 230.

[13] There is considerable controversy about the intended scope of Section 230, as opposed to how it has been interpreted; *see, e.g.,* Solove, Daniel, *Restoring the CDA Section 230 to What It Actually Says*, PRIVACY + SECURITY BLOG (Feb. 4, 2021), https://teachprivacy.com/restoring-the-cda-section-230-to-what-it-actually-says/.

[14] Danielle Keats Citron, *How to Fix Section 230*, Forthcoming. BOSTON UNIVERSITY LAW REVIEW 39 (2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4054906.

[15] U.S. Const. amend. I. *See* Part II.A, *infra*, for a more detailed First Amendment analysis of our proposal.

publishing; indeed, without the circulation, the publication would be of little value.'"[16] Justice Thomas has noted the importance of anonymity to freedom of the press in the Founders 'era: "the historical evidence indicates that Founding-era Americans opposed attempts to require that anonymous authors reveal their identities on the ground that forced disclosure violated the 'freedom of the press.'"[17] Today, there is the same need for anonymity when distributing online content. [18]

Even apart from the constitutional issues, requiring that images carry identification information poses a privacy threat. Web sites often use persistent identifiers—ones not unique to a given session or even a given web site—to track users,[19] often for advertising.[20]

Finally, technical issues exist: there is no current, widely used mechanism for reliably associating online behavior with a specific individual, especially with sufficient reliability to use as evidence in court.[21]

---

[16] *Talley v. California*, 362 U.S. 60, 64 (1960).

[17] *McIntyre v. Ohio Elections Commission*, 514 U.S. 334, 361 (1995) (Thomas, J., concurring).

[18] Goodman, Amy, *Why is Ramsey Orta, Man Who Filmed Police Killing of Eric Garner, the Only One Criminally Charged?*, DEMOCRACY NOW, https://www.democracynow.org/2016/1/12/why_is_ramsey_orta_man_who or Robertson, Adi, *One Tweet Tried to Identify a Cop—Then Five People Were Charged with Felony Harassment*, THE VERGE (Aug. 6, 2020), https://www.theverge.com/2020/8/6/21355999/twitter-cyber-harassment-felony-charges-police-protests-retweet.

[19] Persistent identifiers allow users to be tracked even without knowledge of their names, email addresses, etc. *See* Rich, Jessica, *Keeping Up with the Online Advertising Industry*, FTC BUSINESS BLOG (Apr. 26, 2016), https://www.ftc.gov/news-events/blogs/business-blog/2016/04/keeping-online-advertising-industry ("As the FTC has discussed for years now—see our the 2009 staff report on online behavioral advertising and our 2012 Privacy Report—we regard data as 'personally identifiable,' and thus warranting privacy protections, when it can be *reasonably linked* to a particular person, computer, or device. In many cases, persistent identifiers such as device identifiers, MAC addresses, static IP addresses, or cookies meet this test."). *See also* 16 C.F.R. 312.2(3)(2)(7).

[20] Online advertising generally does not rely on personally identifiable information (PII). Rather, it relies on behavior: web sites visited, search queries, etc. There is a vast literature on this topic; *see, e.g.,* STEVEN M. BELLOVIN, COMMENTS ON PRIVACY 9 (Nov. 2018) ("However, ML algorithms do not need to know someone's identity to invade privacy. The 'My TiVo Thinks I'm Gay' incident is just one example, but in principle, most recommendation algorithms do not need PII."). The cited incident involved a TiVo digital video recorder recommending gay content to someone because of other things he had viewed; *see* Jeffrey Zaslow, *If TiVo Thinks You Are Gay, Here's How to Set It Straight*, WALL STREET JOURNAL (Nov. 26, 2002), https://www.wsj.com/articles/SB1038261936872356908.

[21] Conceivably, if everyone were required to use some sort of government-issued ID

For this reason, the Federal Trade Commission's regulations implementing the Children's Online Privacy Protection Act[22] only apply to websites that are "directed to children"[23] rather than only activities by children: there is no way to know if a child is the actual user of a web site or service. All that said, several jurisdictions are imposing age verification requirements. The United Kingdom's Digital Economy Act[24] provided for the designation of an "age-verification regulator."[25] Louisiana, Utah, and Arkansas have recently enacted laws requiring websites that serve pornography to verify the ages of their users.[26]

We suggest that advanced cryptologic concepts,[27] combined with already proposed changes to Section 230, provide a way forward. Specifically, it is possible to construct privacy-preserving identity credentials with revocable anonymity: the ability for some other party to learn who used such credentials.[28] These credentials could be used to digitally sign uploaded images.[29] In turn, a modification to Section 230 could encourage web sites to take good faith measures to combat intimate

---

every time they logged on to some Internet service, that could provide evidence for actions taken during that session. This is done in China (Haiping Zheng, *Regulating the Internet: China's Law and Practice*, 4 BEIJING L. REV. 37, 39 (2013) ("Internet cafes are required to record every user's identity and online activities… Thus, one who does not take his valid identification with him may not access internet in internet cafes.") but is not the norm in the United States.

[22] Children's Online Privacy Protection Act, 15 U.S.C. § 6501–6506 (1998).

[23] Children's Online Privacy Protection Act, 16 C.F.R. § 312.2.

[24] Digital Economy Act of 2017.

[25] *Id.* § 16.

[26] 2022 La. ALS 440, 2022 La. ACT 440, 2022 La. HB 142 (June 15, 2022); Utah Code Ann. § 78B-3-1002 (2023); Ark. Code Ann. § 4-88-1101 (2023). We note without further analysis that the cryptographic mechanism discussed in this article is capable of conveying an age assertion such as "over 18 years old" in a privacy-preserving fashion.

[27] The technical mechanisms are explained in detail in Part III, *infra*.

[28] Jan Camenisch & Anna Lysyanskaya, *An Efficient System for Non-Transferable Anonymous Credentials with Optional Anonymity Revocation*, *in* ADVANCES IN CRYPTOLOGY — EUROCRYPT 2001, 93 (Birgit Pfitzmann ed., Springer Berlin Heidelberg 2001).

[29] "Digital signatures" are a well-known cryptologic construct, first proposed in 1976 (Whitfield Diffie & Martin E. Hellman, *New Directions in Cryptography*, IT-22 IEEE TRANSACTIONS ON INFORMATION THEORY 644 (Nov. 1976)). How they are employed here is explained in Part III, *infra*. Roughly speaking, they permit a party knowing a secret value to "sign" a document or message, but anyone in the world can verify the signature without knowing the secret.

image abuse by mechanisms such as these credentials. Such changes to Section 230 have already been proposed.[30]

As with any mechanism for criminal sanctions, our proposal is not a panacea. At best, criminal law acts as a deterrent; with nonconsensual pornography, as with many other crimes, the harm done is irremediable.[31] Still, such activity is illegal in most of the United States;[32] enhancing the ability to find and prosecute offenders should help deter violations. It is well known that crime is deterred more by certainty of punishment than by its severity.[33]

We do not claim to have complete answers. As discussed later in this article and especially in Part V, *infra*, there are a number of difficult economic and social issues that must be addressed in future work. But we think that this is a promising avenue to explore.

The technical details of our system are complex; we defer discussion of it to Part III, *infra*. For now, let it suffice to say that for users to upload images to participating web sites, the web sites would need to run some special software and users would need to register with an *Identity Provider Service* and use a suitable web browser in order to perform uploads.

Our proposal raises a number of difficult doctrinal questions, especially involving the First Amendment. These include:

- Does requiring or incentivizing web sites to to implement this credential system amount to compelled speech, in violation of the First Amendment? If it is compelled speech, a somewhat different legal analysis is necessary.
- Is the constitutional right to anonymous speech chilled by the possibility of deanonymization?
- Is the constitutional right to free speech disproportionately burdened by the prior registration requirement?
- What level of scrutiny should be applied to these First Amendment issues?

---

[30] *See, e.g.,* Danielle Keats Citron & Wittes, Benjamin, *The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity*, 86 FORDHAM LAW REVIEW 401, 419 (Nov. 2017) ("No provider or user of an interactive computer service that takes reasonable steps to prevent or address unlawful uses of its services shall be treated as the publisher or speaker of any information provided by another information content provider in any action arising out of the publication of content provided by that information content provider.")

[31] *See, e.g.,* Danielle Keats Citron & Franks, Mary Anne, *Criminalizing Revenge Porn*, 49 WAKE FOREST LAW REVIEW 345 (2014).

[32] *See* note 9, *supra*.

[33] *See, e.g.,* National Institute of Justice, "Five Facts About Deterrence", May 2016, *available at* https://www.ojp.gov/pdffiles1/nij/247350.pdf.

- What statutory or regulatory changes are needed to implement this scheme?

II. Legal Background and Analysis

Any solution to the problem of non-consensual pornography must contend with a variety of First Amendment[34] issues. We are suggesting, at the least, that the government strongly encourage use of certain software; furthermore, the purpose of this software is to revoke anonymity. Both parts are constitutionally problematic, though we think in this case justifiable. There are other issues involving the necessary modifications to Section 230. Here, we briefly review some of the major legal issues.

A.First Amendment Issues

If the code we suggest be used is expressive, and hence more strongly protected by the First Amendment, sites arguably could not be forced to run it unless the entire system passes strict scrutiny. To be sure, this is likely commercial speech and hence less strongly protected.[35] Still, even commercial speech has some First Amendment protections.[36] We therefore start our analysis by analyzing if the software is speech.

Our first issue is that we seek to prescribe functionality of computer code run by assorted websites, and perhaps to mandate that these companies write—speak—certain code. While the Supreme Court has never ruled directly on the issue, several lower courts have held that computer code can be speech, and hence be protected by the First Amendment. The first case to address the issue squarely was *Bernstein v. United States Department of State*.[37] Bernstein wished to publicly post—and hence export—source code

---

[34] U.S. Const. amend I.

[35] *Cent. Hudson Gas & Elec. Corp. v. Public Serv. Comm'n*, 447 U.S. 557, 562–63 (1980) ("The Constitution therefore accords a lesser protection to commercial speech than to other constitutionally guaranteed expression.")

[36] *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 557 (2011) ("Speech in aid of pharmaceutical marketing, however, is a form of expression protected by the Free Speech Clause of the First Amendment.")

[37] *Bernstein v. United States Dep't. of State*, 922 F. Supp. 1426 (N.D. California 1996). The 9th Circuit upheld this ruling, *Bernstein v. United States DOJ*, 176 F.3d 1132 (9th Cir. 1999), in a ruling that was withdrawn when the court agreed to rehear the case *en banc*. That, in turn, was mooted when the United States changed its export rules for cryptographic software; *see Bernstein v. DOC*, 2004 U.S. Dist. LEXIS 6672, 2004 WL 838163, n. 2 (N.D. Cal 2004) ("before the rehearing could take place, defendants announced plans to make additional changes to the EAR. In January 2000, defendants added 15 C.F.R. section 740.13(e) to the Federal Register, which allows the DOC to exempt 'publicly available' encryption source code from license requirements.")

for some cryptographic algorithms, in apparent violation of the export control rules of the time.[38] A crucial issue was whether computer code was purely functional as opposed to expressive; the court concluded that in this instance, it was both, and hence was protected by the First Amendment.

There was an important limitation in the *Bernstein* court's holding: what was the intended purpose of the code? "We emphasize the narrowness of our First Amendment holding. We do not hold that all software is expressive. Much of it surely is not. Nor need we resolve whether the challenged regulations constitute content-based restrictions, subject to the strictest constitutional scrutiny, or whether they are, instead, content-neutral restrictions meriting less exacting scrutiny. We hold merely that because the prepublication licensing regime challenged here applies directly to scientific expression, vests boundless discretion in government officials, and lacks adequate procedural safeguards, it constitutes an impermissible prior restraint on speech."[39]

Similarly, the purpose of the code was crucial in a case where a court held that the code in question was not speech, *Universal City Studios v. Corley.*[40] Eric Corley, the publisher of *2600 Magazine*, ran an article, including source code, on decrypting DVDs. Universal City Studios sued, asserting that this violated the anti-circumvention provision Digital Millennium Copyright Act.[41] Corley asserted that he had a First Amendment right to publish the program at issue. The Second Circuit disagreed: "The Appellants 'argument fails to recognize that the target of the posting provisions of the injunction—DeCSS—has both a nonspeech and a speech component, and that the DMCA, as applied to the Appellants, and the posting prohibition of the injunction target only the nonspeech component. Neither the DMCA nor the posting prohibition is concerned with whatever capacity DeCSS might have for conveying information to a

[38] International Traffic in Arms Regulations (ITAR), 22 C.F.R. §§ 120-30 (1994). Bernstein's suit was one piece of a larger movement to democratize access to cryptography; *see generally* STEVEN LEVY, CRYPTO: HOW THE CODE REBELS BEAT THE GOVERNMENT—SAVING PRIVACY IN THE DIGITAL AGE (Viking 2001). In an earlier suit, *Karn v. United States Dep't of State*, 925 F. Supp. 1 (1996)., the court declined to rule on whether computer programs constituted speech, *id.* n. 19 ("The Court makes no ruling as to whether source codes, without the comments, fall within the protection of the First Amendment.").

[39] *Bernstein v. United States DOJ*, 176 F.3d 1132, 1145 (9th Cir. 1999).

[40] *273 F.3d 429 (2nd Cir. 2001).*

[41] 17 U.S.C. 1201. The relevant passage, 17 U.S.C. 1201(a)(2) (A), provides that "No person shall manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof, that… is primarily designed or produced for the purpose of circumventing a technological measure that effectively controls access to a work protected under this title."

human being, and that capacity, as previously explained, is what arguably creates a speech component of the decryption code."[42] In other words, while the expressive content of computer code is constitutionally protected, its functional component might not be.

The Sixth Circuit also noted the dual nature of computer code.[43] While holding that "computer source code is an expressive means for the exchange of information and ideas about computer programming,"[44] it also noted that the "functional capabilities of source code, and particularly those of encryption source code, should be considered when analyzing the governmental interest in regulating the exchange of this form of speech."[45]

We must, therefore, consider several separate issues. First, of course, is whether code can be considered as speech. The *Bernstein*, *Universal City Studios*, and *Junger* courts all agreed that code can be speech. The next question is the purpose of the code. The *Bernstein* court held that the purpose of that particular publication request was scientific: "We hold merely that because the prepublication licensing regime challenged here applies directly to scientific expression, vests boundless discretion in government officials, and lacks adequate procedural safeguards, it constitutes an impermissible prior restraint on speech."[46] The second, per *Universal City Studios*, is the balance between the functional and expressive components. Finally, per *Junger*, the balance and purpose must be considered when deciding what level of scrutiny to apply.[47]

Given that, deanonymization code mandated for, say, a political web site dedicated to displaying pictures of police brutality would be considered compelled speech violating the expressive intent of the site (and thus unconstitutional unless it survived strict scrutiny), while the same requirement as applied to a pornography site would require a lesser level. It is not that the computer code in the latter case is less expressive; however, the purpose of the code, preventing full anonymity, could be balanced against the government's interests in preventing non-consensual pornography.

Government-compelled speech is generally unconstitutional. In the classic case *West Virginia State Board of Education v. Barnette*,[48] in which

---

[42] *273 F.3d 429, 454 (2001).*

[43] *Junger v. Daley*, 209 F.3d 481 (6th Cir. 2000). This case was also an outgrowth of the movement described by Levy, *supra* note 37.

[44] *Junger*, 209 F.3d at 485.

[45] *Id.*

[46] *Bernstein v. United States DOJ*, 176 F.3d 1132, 1145 (9th Cir. 1999).

[47] *Junger v. Daley*, 209 F.3d 481, 485 (6th Cir. 2000).

[48] *319 U.S. 624 (1943).*

a West Virginia law required students to salute the American flag was struck down, the Court wrote "If there is any fixed star in our constitutional constellation, it is that no official, high or petty, can prescribe what shall be orthodox in politics, nationalism, religion, or other matters of opinion or force citizens to confess by word or act their faith therein."[49] This is a principle that the Court has consistently upheld: "Some of this Court's leading First Amendment precedents have established the principle that freedom of speech prohibits the government from telling people what they must say."[50] To the extent that computer code is pure speech, it cannot be restricted unless there are some sort of exceptional circumstances.

With the exception of *Universal City Studios*, though, these anonymity and code-as-speech cases are not quite apposite. *McIntyre* and *Talley* invalidated overly broad statutes, in cases involving political speech, i.e., implicating core First Amendment values. Similarly, Bernstein and Junger were trying to make political points with their attempts to get export licenses and hence with their suits. Karn's case was even more political: he was denied an export license for cryptographic source code when he had been granted a license for a book containing the exact same code.[51] In *Universal City Studios*, though, the apparent purpose of the code was a violation of the anti-circumvention provisions of the DMCA. That is, its functional element was a violation of the law. The court held that "[t]he Government's interest in preventing unauthorized access to encrypted copyrighted material is unquestionably substantial, and the regulation of DeCSS by the posting prohibition plainly serves that interest. Moreover, that interest is unrelated to the suppression of free expression."[52] Computer code whose *functional* purpose is illegal is not protected by the First Amendment. This is not quite our situation, in that we are not trying to ban code that violates the law; however, the purpose of the code that we propose is to deter illegal behavior.[53] That is, the intent of the code is different.

The situation here is thus more nuanced than in, *e.g., Barnette*. *Barnette* dealt with core First Amendment interests of individuals, freedom of speech and religion. Even if code is speech, our scheme implicates compelled commercial speech, an entirely different situation. In *Zauderer v. Office of Disciplinary Counsel of the Supreme Court of Ohio*,[54] the Court held that

---

[49] *Id.* at 642.

[50] *Rumsfeld v. Forum for Academic & Institutional Rights, Inc.*, 547 U.S. 45, 61 (2006).

[51] *Karn v. United States Dep't of State*, 925 F. Supp 1, 3 (1996).

[52] *Universal City Studios v. Corley*, 273 F.3d 429, 454 (2001).

[53] As noted *supra*, non-consensual pornography is illegal in most states; *see* Citron & Franks, Mary Anne, *supra* note 31.

[54] 471 U.S. 626 (1985).

compelling truthful disclosures in advertising was permissible: "Because the extension of First Amendment protection to commercial speech is justified principally by the value to consumers of the information such speech provides… appellant's constitutionally protected interest in *not* providing any particular factual information in his advertising is minimal. (Internal citation omitted)"[55] To be sure, this decision was not a blanket endorsement on compelled commercial speech,[56] but it did leave room for reasonable regulation.[57]

A few years before these cases, the Court laid out its framework for regulation of commercial speech in *Central Hudson Gas and Electric v. Public Service Commission*.[58] It wrote that:

> In commercial speech cases, then, a four-part analysis has developed. At the outset, we must determine whether the expression is protected by the First Amendment. For commercial speech to come within that provision, it at least must concern lawful activity and not be misleading. Next, we ask whether the asserted governmental interest is substantial. If both inquiries yield positive answers, we must determine whether the regulation directly advances the governmental interest asserted, and whether it is not more extensive than is necessary to serve that interest.[59]

We are dealing here with compelled speech, a somewhat different issue, so the issues of lawful activity and truthfulness are not relevant. The other points, though, are important: is the expression—running the necessary code—covered by the First Amendment, does this directly advance a government interest, and is the regulation more extensive than necessary?

Zauderer was about deceptive speech. The First Circuit noted, though, that "we have found no cases limiting *Zauderer* in such a way,"[60] a view echoed by the Second Circuit in *N.Y. State Rest. Ass'n v. N.Y. City Bd. of Health (NYSRA)*.[61] Rather, courts have upheld compelled speech if it is commercial, factual, and not controversial. For example, the Second Circuit held that "Required disclosure of accurate, factual commercial information

---

[55] *Id.* at 651.

[56] *Id.* ("We recognize that unjustified or unduly burdensome disclosure requirements might offend the First Amendment by chilling protected commercial speech.")

[57] The *Zauderer* court did not specify the level of scrutiny to be applied, but indicated that this sort of regulation, requiring truthful advertising, did not require strict scrutiny. *Id.* n.14.

[58] *447 U.S. 557 (1980).*

[59] *Id.* at 566.

[60] *Pharm. Care Mgmt. Ass'n v. Rowe*, 429 F.3d 294, 310 n.8 (1st Cir. 2005).

[61] 556 F.3d 114, 133 (2nd Cir. 2009).

presents little risk that the state is forcing speakers to adopt disagreeable state-sanctioned positions, suppressing dissent, confounding the speaker's attempts to participate in self-governance, or interfering with an individual's right to define and express his or her own personality."[62] Our scheme is, arguably, linked to dissent and self-governance, which is why we suggest a higher standard of scrutiny than the rational basis standard applied in *NYSRA*.[63]

There is one more point to consider. Some online services are explicitly designed or intended to protect user privacy and anonymity. Protonmail, a Swiss email provider, is available via Tor;[64] they cite anonymity as one of their reasons for offering their service that way: "Tor also makes your connections to Proton anonymous. We will not be able to see the true IP address of your connection to Proton."[65] Another email provider, Lavabit, claims (though without giving details) that it keeps minimal metadata about communications to protect user privacy: "Minimized metadata: Who you communicate with is as private as what you say."[66] In fact, the company once fought a court order to make communications metadata readily available to the government.[67] The company declined to cooperate with the court order "because 'Lavabit did not want to "defeat [its] own system."'"[68] Arguably, anonymity code used by such companies is expressive and hence speech, since privacy is part of these companies '*raison d'être*.

As we have shown, the issue of whether computer code constitutes speech is complex and not easily answered. If code is not speech, there is no First Amendment issue. Let us assume, then, that it is speech. The problem of non-consensual pornography is serious; there is thus a substantial

---

[62] *Nat'l Elec. Mfrs. Ass'n v. Sorrell*, 272 2nd Cir. 104, 114 (2001).

[63] *NYSRA*, 556 F.3d at 134.

[64] Roger Dingledine et al., *Tor: The Second-Generation Onion Router*, Proceedings of the 13th USENIX Security Symposium (Aug. 2004).

[65] *Access all Proton services with the Tor anonymity network*, https://proton.me/tor (last visited July 14, 2023).

[66] *Security*, https://lavabit.com/security.html, (last visited July 14, 2023).

[67] *United States v. Lavabit (In re Under Seal)*, 749 F.3d 276, 280–281 (4th Cir.) ("On June 28, 2013, the Government sought and obtained an order ('the Pen/Trap Order') from a magistrate judge authorizing the placement of a pen register and trace-and-trap device on Lavabit's system. This 'pen/trap' device is intended to allow the Government to collect certain information, on a real-time basis, related to the specific investigatory target's Lavabit email account. In accordance with the Pen/Trap Statute, 18 U.S.C. §§ 3121-27, the Pen/Trap Order permitted the Government to "capture all non-content dialing, routing, addressing, and signaling information . . . sent from or sent to" the target's account. In other words, the Pen/Trap Order authorized the Government to collect metadata relating to the target's account, but did not allow the capture of the contents of the target's emails.")

[68] *Id.* at 281.

government interest in solving it. We thus turn to the question of extensiveness of the government intrusion.

As described in Part V.A, *infra*, other means of tracing who initially upload non-consensual pornography are not likely to work. The scheme we propose is considerably stronger, and, though not foolproof, is much more likely to provide the necessary information. We thus conclude that it satisfies the last prong of the *Central Hudson* test.[69]

Another, and perhaps more serious, point of conflict with the First Amendment is that the scheme we propose is designed to limit anonymity.[70] The Supreme Court, though, has long held that anonymous speech is constitutionally protected. In *Talley v. California*,[71] the Court invalidated a statute that prohibited anonymous leafleting. The city of Los Angeles had an ordinance that required all leaflets to contain the true names and addresses of the author and distributors. In the case at issue, the leaflets in question advocated a boycott of some stores on civil rights grounds: the businesses allegedly "carried products of 'manufacturers who will not offer equal employment opportunities to Negroes, Mexicans, and Orientals.'"[72] This was clearly political speech, the type of speech most strongly protected by the First Amendment. The Court noted that "[t]here can be no doubt that such an identification requirement would tend to restrict freedom to distribute information and thereby freedom of expression."[73] They went on to say that "[a]nonymous pamphlets, leaflets, brochures and even books have played an important role in the progress of mankind."[74] Significantly, the ruling turned in part on the lack of a suitable purpose in the challenged statute: its language did not refer to fraud, libel, etc. Had the statute contained an explicit limitation of purpose, it might have survived.[75] Furthermore, as Justice Harlan noted, "In the absence of a more substantial showing as to Los Angeles 'actual experience with the distribution of

---

[69] *Cent. Hudson Gas & Elec. Corp. v. Public Serv. Comm'n*, 447 U.S. 557, 566 (1980).

[70] For a thorough look at anonymity and the First Amendment, *see generally* JEFF KOSSEFF, THE UNITED STATES OF ANONYMOUS (Cornell University Press 2022). *See also* Leeza Arbatman & John Villasenor, *Anonymous Expression and "Unmasking" in Civil and Criminal Proceedings*, 23 MINN. J.L. SCI. & TECH. 77 (2022).

[71] 362 U.S. 60 (1960).

[72] *Id.* at 61.

[73] *Id.* at 64.

[74] *362 U.S. 60, 64 (1960).*

[75] *Id.* ("Counsel has urged that this ordinance is aimed at providing a way to identify those responsible for fraud, false advertising and libel. Yet the ordinance is in no manner so limited, nor have we been referred to any legislative history indicating such a purpose. Therefore we do not pass on the validity of an ordinance limited to prevent these or any other supposed evils.").

obnoxious handbills, such a generality is for me too remote to furnish a constitutionally acceptable justification for the deterrent effect on free speech which this all-embracing ordinance is likely to have."[76] Here, the evidence for the existence of and harm caused by non-consensual pornography is overwhelming.

The Court reiterated that point in *McIntyre v. Ohio Elections Commission*,[77] where it struck down a law prohibiting anonymous campaign literature. This ordinance differed somewhat from the statute invalidated in *Talley*; this applied only to campaign literature. The court struck it down nevertheless: "The Ohio statute likewise contains no language limiting its application to fraudulent, false, or libelous statements; to the extent, therefore, that Ohio seeks to justify §3599.09(A) as a means to prevent the dissemination of untruths, its defense must fail for the same reason given in *Talley*."[78] The Court went on to say that "consequently, we are not faced with an ordinary election restriction; this case 'involves a limitation on political expression subject to exacting scrutiny.'"[79] In other words, the Ohio statute burdened political speech and was not narrowly tailored.

The Supreme Court has been willing to uphold identity disclosure requirements, even in a political context, when a compelling state interest was shown. In *Citizens United v. FEC,*[80] a non-profit organization wanted to distribute a documentary film about Hillary Clinton, in violation of the Bipartisan Campaign Reform Act.[81] While the case is best known because the Court used it to strike down limits on corporate campaign activities, the Court explicitly upheld a portion of the law requiring political ads to identify their source. The Courte wrote "The disclaimers required by §311 'provid[e] the electorate with information, 'and 'insure [sic] that the voters are fully informed 'about the person or group who is speaking… ('Identification of the source of advertising may be required as a means of disclosure, so that the people will be able to evaluate the arguments to which they are being subjected'). At the very least, the disclaimers avoid confusion by making clear that the ads are not funded by a candidate or political party."[82] The Court also noted that, "Citizens United finally claims

[76] *Talley*, 362 U.S. at 66–67 (Harlan, J., concurring).

[77] *McIntyre v. Ohio Elections Commission*, 514 U.S. 334 (1995).

[78] *Id.* at 344.

[79] *Id.* at 345. More likely, the Court meant "strict scrutiny", since the statute implicated core First Amendment values.

[80] 558 U.S. 310 (2010).

[81] 52 U.S.C. § 30101.

[82] *Citizens United*, 558 U.S. at 368.

that disclosure requirements can chill donations by exposing donors to retaliation, but offers no evidence that its members face the type of threats, harassment, or reprisals that might make §201 unconstitutional as applied."[83] Significantly, the Court subjected the disclosure requirements "to 'exacting scrutiny, 'which requires a 'substantial relation 'between the disclosure requirement and a 'sufficiently important 'governmental interest."[84] By contrast, the core rulings on political speech relied on strict scrutiny.[85]

Other courts have held that recording of identities can be legitimate where the purpose of the statute is to protect children against commercial sexual abuse.[86] In a pair of statutes, Congress has required that producers of imagery[87] involving actual or simulated sexual contact could be required to record the names of the performers, and to make those records available to government investigators.[88] Rejecting a claim that these statutes violated the constitutional right to anonymous speech, the Sixth Circuit wrote that "Section 2257, however, does none of these things [affecting a broad spectrum of speech, hindering an historically significant mode of communication and destroying anonymous and spontaneous advocacy by making the registration records open to the public at large]: It affects only a narrow category of speech and does so for the limited purpose of preventing

[83] *Id.* at 370.

[84] *Id.* at 366–367.

[85] *Id.* at 340 ("Laws that burden political speech are 'subject to strict scrutiny,' which requires the Government to prove that the restriction 'furthers a compelling interest and is narrowly tailored to achieve that interest.'") (Internal citations omitted.)

[86] *Connection Distrib. Co. v. Holder*, 557 F.3d 321, 325 (6th Cir. 2009) ("In 1986, the Attorney General's Commission on Pornography determined that, although efforts to eradicate child pornography had 'drastically curtailed its public presence,' they 'ha[d] not ended the problem.'… Prompted by the Commission's report and recommendations, Congress in 1988 enacted the Child Protection and Obscenity Enforcement Act… Section 7513(a) of the Act, known by its codified section number as § 2257, attempted to address this problem by adding a reporting and verification requirement to the existing laws designed to prevent child pornography.")

[87] "Produces" is defined in 18 U.S.C. § 2257(h)(2), note 82, *infra*. The definition specifically excludes web sites that host such material (18 U.S.C. § 2257(h)(2)(B)). *But see* 28 C.F.R. § 75.1(c)(2), defining "secondary producer", and *Free Speech Coalition v. Gonzales*, 406 F. Supp. 1196, 1204 (D. Col.), enjoining enforcement of much of the secondary producer regulation ("the statute and regulations may not be enforced as to secondary producers who are not involved in any activity that involves 'hiring, contracting for managing, or otherwise arranging for the participation of the performers depicted.'").

[88] Child Protection and Obscenity Enforcement Act of 1988, 18 U.S.C. § 2257; Adam Walsh Child Protection and Safety Act of 2006, 18 U.S.C. § 2257A.

speech (child pornography) that the First Amendment does not protect."[89] The Third Circuit ruled similarly.[90] In other words, the right to anonymity of performers of adult content can constitutionally be limited by statute.

It is important to realize, though, that photographs can serve core First Amendment purposes by documenting official misconduct. Indeed, many of the Black Lives Matter protests of the last several years have been triggered by photographs or videos of apparent police misbehavior.[91] In today's climate, some people will only publish their evidence if shielded by anonymity; rightly or wrongly, they fear reprisal by other police officers.[92]

Our proposal is for a narrowly tailored requirement, that deanonymization be possible only for content that judges have found to have probable cause to be illegal under the statute we discuss in Part V, *infra*. Furthermore, as discussed in Part IV, *infra*, we propose several structural mechanisms to prevent abuse. In that section, we also discuss the legal tests that should be applied.

While the Supreme Court has never ruled on the issue, a number of lower and state courts have considered the question of deanonymizing online activity. As Professor Jeff Kosseff notes, courts are generally more

---

[89] *Connection Distrib. Co.*, 557 F.3d at 333. *But see id.* at 347, Kennedy, dissenting ("Registration requirements have been recognized to have a significant chilling effect on speech because they force those who would speak anonymously 'to forgo their right'.")

[90] *Free Speech Coal. v. AG United States*, 787 F.3d 142, 158 (3rd Cir. 2015) ("Suffice it to say, however, that the Government's interest in ensuring that minors are not sexually exploited is advanced when completely anonymous sexual participants whom we have no reason to believe are over 18 must verify their ages before appearing in sexually explicit materials.")

[91] *See, e.g.,* the many incidents recounted in Marc Freeman, *A Mom Got Arrested for Videotaping Cops in Public. Were Her Rights Violated?*, SOUTH FLORIDA SUN SENTINEL (Dec. 14, 2020), https://www.sun-sentinel.com/local/palm-beach/fl-ne-police-videorecording-public-lawsuit-appeal-ss-prem-20201214-tbomhpwnyjbizjhfhnmgnr7ewi-story.html. *See also* Howard M Wasserman, *Orwell's Vision: Video and the Future of Civil Rights Enforcement*, 68 MD. L. REV. 600 (HeinOnline 2008).

[92] Stories of apparent police reprisal for First Amendment-protected activities are legion; *see, e.g.,* Mikki Kendall, *The Police Can't Police Themselves. And Now the Public is Too Scared to Cooperate with Them.*, WASHINGTON POST (Apr. 10, 2015), https://www.washingtonpost.com/posteverything/wp/2015/04/10/the-police-cant-police-themselves-and-now-the-public-is-too-scared-to-cooperate-with-them/ or Laura Vozzella & Gregory S. Schneider, *Virginia Sen. L. Louise Lucas Charged with Felonies over Portsmouth's Confederate Monument Protest*, (Aug. 17, 2020), https://www.washingtonpost.com/local/virginia-politics/virginia-sen-l-louise-lucas-charged-with-felonies-over-portsmouth-confederate-monument-protest/2020/08/17/84fd4bf6-e0c8-11ea-8181-606e603bb1c4_story.html.

willing to do so in case of alleged criminal activity;[93] however, our scheme must provide mechanisms useful in civil suits, since some states provide for a private right of action as well for non-consensual pornography.[94]

We must also consider the problem of "chilling effects":[95] that people will self-censor because of the fear of what might happen. Schauer defines the concept this way: "the very essence of a chilling effect is an act of deterrence."[96] He primarily analyzes it from a perspective of the law: people will refrain from saying something because they fear that they might run afoul of a statute. The concept is a "major substantive component of First Amendment adjudication,"[97] so much so that by 1967 Justice Harlan called the doctrine "ubiquitous."[98]

Penney looks at it more broadly: a chilling effect has to be seen in social terms, as well as legal ones: people may be afraid of violating social norms and not just laws: "chilling effects predominantly involve not just a deterrent effect, but a shaping effect—people speaking, acting, or doing, in a way that conforms to, or is in compliance with, a perceived social norm, not simply self-censoring to avoid a legal harm."[99] He quotes Quentin Skinner[100] as noting that 17th century writers understood this: "arguing that our mere awareness of living under an arbitrary power—a power capable of interfering with our activities without having to consider our interests—serves in itself to limit our liberty. Knowing that we are free to do or forbear only because someone else has chosen not to stop us is what reduces us to servitude."[101] Will the possibility of deanonymization chill legal anonymous speech?

---

[93] KOSSEFF, *supra* note 67, at 158 ("courts presiding over criminal cases and grand jury investigations tend to place less emphasis on protecting anonymity").

[94] *See, e.g.,* N.Y. Civ. Rights Law § 52-b; 13 V.S.A. § 2606 (e)(1); Cal. Civ. Code § 1708.85.

[95] The classic paper on chilling effects in a First Amendment context is Frederick Schauer, *Fear, Risk, and the First Amendment:Unraveling the Chilling Effect*, 58 B.U. L. REV. 685 (1978). For a broader look and copious references to the history of the concept, *see generally* Jonathon W. Penney, *Understanding Chilling Effects*, 106 MINN L. REV. 1451 (2022).

[96] Schauer, *supra* note 92, at 689.

[97] *Id.* at 685.

[98] *Zwickler v. Koota*, 389 U.S. 241, 256 n.2 (1967) (Harlan, J., concurring), quoted in Schauer, *supra* note 89 at 685 n.3.

[99] Penney, *supra* note 92, at 1455.

[100] *Id.* at 1463 n.43.

[101] Quentin Skinner, *A Third Concept of Liberty*, 24 LONDON REVIEW OF BOOKS (Apr. 4, 2002), https://www.lrb.co.uk/the-paper/v24/n07/quentin-skinner/a-third-concept-of-liberty.

The Supreme Court has long recognized the possibility of non-legal effects chilling First Amendment-protected activity. In *NAACP v. Alabama,* the Court blocked a statutory requirement that the NAACP disclose a list of its members to the state: "Petitioner has made an uncontroverted showing that on past occasions revelation of the identity of its rank-and-file members has exposed these members to economic reprisal, loss of employment, threat of physical coercion, and other manifestations of public hostility."[102]

In our scheme, chilling effects can arise because of the possibility that the identity of the image poster will be disclosed. Per Penney, the mere possibility that this can happen can cause a chilling effect, even if it never actually happens absent actual legal cause: "For example, a growing body of research in social-psychology has documented what has been called a 'watching eye 'effect, wherein artificial surveillance cues—like simply a set of 'watchful 'human eyes in the presence of participants—can have a chilling effect on their behavior. That is, the awareness of surveillance—even where participants *know* it is artificial and nobody is *actually* watching—promotes socially conforming or compliant behavior in a wide range of contexts… Interestingly, research shows that even where the 'watching eye 'is clearly artificial (e.g., the 'gaze 'deployed is simply a photo or image of an eye) these effects on behavior remain."[103]

There is thus no doubt that our scheme will have some chilling effect. Against this, we must balance the chilling effect of non-consensual pornography. Eaton *et al.* found that a significant number of people were simply threatened with publication of intimate images, with no images actually being posted.[104] Lenhart *et al.* noted that "Even if the images are never actually posted publically [sic], the perpetrator may use threats to post such images as a method of controlling or intimidating the victim."[105] In other words, the chance that intimate images will be posted publicly can have a chilling effect on what otherwise might be consensual sexual behavior.

We must also examine the costs of our scheme, and in particular the extent to which these costs might differentially inhibit First Amendment-protected activity.[106] There is no doubt that the scheme we propose will

[102] 357 U.S. 449, 462 (1958).

[103] Penney, *supra* note 92, at 1483–1484.

[104] Eaton et al., *supra* note 5, at 11.

[105] Lenhart et al., *supra* note 7, at 4.

[106] Many courts have examined the constitutionality of laws against nonconsensual pornography. A partial list is given in *State v. Katz*, 179 N.E.3d 431, 450 (Ind. 2022) ("And none of these statutes have ultimately been struck down as unconstitutional.") The Illinois

cause some burden on the ability of some people to speak on the Internet. Those who do not have the necessary credentials from an Identity Provider will not be able to post images to many web sites. The deeper issue, though, is not so much the burden—as will be discussed, there have always been economic burdens to the exercise of freedom of speech and the press—but rather the creation of a government-encouraged restriction that disparately burdens some people. That is, obtaining the necessary credentials will pose a considerable hurdle to many, especially the poor, the underprivileged, and those who live in remote areas—they may not have the necessary government-issued documents to prove their identity, and the cost of an online service, which is minimal to some, may prove an insurmountable obstacle to others.[107]

The lack of money, of course, has long been an obstacle to freedom of speech and of the press. The Supreme Court noted in 1974 that "[t]he obvious solution, which was available to dissidents at an earlier time when entry into publishing was relatively inexpensive, today would be to have additional newspapers. But the same economic factors which have caused the disappearance of vast numbers of metropolitan newspapers, have made entry into the marketplace of ideas served by the print media almost impossible."[108]

The Internet was supposed to change that. "It is no exaggeration to conclude that the Internet has achieved, and continues to achieve, the most participatory marketplace of mass speech that this country—and indeed the world—has yet seen. The plaintiffs in these actions correctly describe the 'democratizing 'effects of Internet communication: individual citizens of limited means can speak to a worldwide audience on issues of concern to them. Federalists and Anti-Federalists may debate the structure of their government nightly, but these debates occur in newsgroups or chat rooms rather than in pamphlets. Modern-day Luthers still post their theses, but to electronic bulletin boards rather than the door of the Wittenberg Schlosskirche."[109]

It hasn't worked out that way—the Internet, like mass media, is dominated by a few large companies whose policies are often antithetical to

---

case, *People v. Austin*, 153 N.E.3d 439 (Ill.), was appealed to the U.S. Supreme Court, which denied certiorari, *Austin v. Illinois*, 141 S. Ct. 233 (2020).

[107] This issue, and some possible solutions, are discussed in more detail in Part V.B, *infra*.

[108] *Miami Herald Pub. Co., Div. of Knight Newspapers, Inc. v. Tornillo*, 418 U.S. 241, 251, 94 (1974).

[109] *ACLU v. Reno*, 929 F. Supp 824, 881 (Third Circuit, E.D. Pennsylvania: District Court 1996).

free speech. Verizon's former Acceptable Use Policy said that users may not use their Internet service to "to damage the name or reputation of Verizon, its parent, affiliates and subsidiaries."[110] Their wireless service once barred text messages in favor of abortion rights.[111] Facebook bars "[u]sing a name that is not the authentic name you go by in everyday life."[112] Twitter[113] has blocked a political campaign ad that focused on abortion rights.[114] None of these restrictions would be permissible if imposed by a government, but private companies can, of course, do as they wish. In other words, there is already a considerable burden to free speech on the Internet.

However, we must also balance the burden we would impose—which, as noted, will fall disproportionately on disadvantaged groups—with the harm done by non-consensual pornography to such groups. Studies have shown that members of such communities are in fact disproportionately victimized. Lenhart *et al.* found that Black people have had NCP photos posted significantly more often than whites, and that those making less than $50,000 per year suffer significantly more than those making $75,000.[115] We conclude, then, that though certain underprivileged groups would be more heavily burdened by our proposal, they are also more likely to be protected by it. This suggests that perhaps the differential impact is acceptable.

Assuming, as suggested, that code is speech. There then remains the question of the level of scrutiny to be applied, given the infringement on First Amendment rights. Such rights are, of course, not absolute. The possible choices here would seem to be intermediate scrutiny, exacting

---

[110] Archived copy at *Verizon Online—Terms of Service*, https://web.archive.org/web/20070820030244/http://www.verizon.net/policies/vzcom/tos_popup.asp (last visited Jun. 18, 2023) Acceptable Use Policy, § 3(j).

[111] Adam Liptak, *Verizon Blocks Messages of Abortion Rights Group*, NEW YORK TIMES (Sep. 27, 2007), https://www.nytimes.com/2007/09/27/us/27verizon.html.

[112] *Account Integrity and Authentic Identity*, Facebook, https://transparency.fb.com/policies/community-standards/account-integrity-and-authentic-identity/ (June 18, 2023).

[113] On July 23, 2023, Elon Musk began renaming Twitter as X. *See* Ryan Mac & Tiffany Hsu, *From Twitter to X: Elon Musk Begins Erasing an Iconic Internet Brand*, NEW YORK TIMES (Jul. 24, 2023), https://www.nytimes.com/2023/07/24/technology/twitter-x-elon-musk.html.

[114] Erik Uebelacker, *Twitter Blocks Democrat's Abortion Rights Campaign Ad*, DAILY BEAST (Jun. 15, 2023), https://www.thedailybeast.com/twitter-blocks-democrats-abortion-rights-campaign-ad.

[115] Lenhart et al., *supra* note 7, at 6.

scrutiny, and strict scrutiny. As we shall see, exacting scrutiny seems to be the right choice.

Under strict scrutiny, a law restricting speech is invalid "unless it is justified by a compelling government interest and is narrowly drawn to serve that interest."[116] Our argument, though, rests on balancing one set of rights under the First Amendment against the right to sexual privacy. While protecting individuals 'sexual privacy is indeed a compelling government interest, this sort of balancing of different constitutional rights falls more naturally under the exacting scrutiny test.

There is some confusion on the difference between "strict scrutiny" and "exacting scrutiny." Professor R. George Wright warns against seeing it as a compromise level between strict scrutiny and intermediate scrutiny.[117] He explains that "[m]ost typical exacting scrutiny formulations may require the governmental regulation at issue to bear something like a 'substantial relation 'to a sufficiently important governmental interest. Balancing and proportionality plainly inhere in the idea of a sufficiently important government interest, given that the sufficiency of the government interest must be with respect to some corresponding burden imposed upon constitutional rights."[118] Put another way, "[i]n Justice Breyer's view, proportionalism may be called for 'when a statute restricts one constitutionally protected interest in order to further some comparably important interest.'"[119] By contrast, few statutes survive strict scrutiny.[120]

Intermediate scrutiny is too weak a standard. The *Universal City Studios* court upheld the trial court's determination that the DMCA prohibition of anti-circumvention devices was content-neutral and that intermediate scrutiny was thus appropriate.[121] Our scheme is not content-neutral, in that revocation of anonymity is intended solely for non-consensual pornography.

We thus opt for exacting scrutiny as the standard of evaluation. Our scheme is designed to meet precisely these balancing and proportionality tests, weighing the right to publish anonymously against the right to

---

[116] *Brown v. Entm't Merchs. Ass'n*, 564 U.S. 786, 799 (2011).

[117] R. George Wright, *A Hard Look at Exacting Scrutiny*, 85 UMKC LAW REVIEW 207, Part II.

[118] *Id.* at 209–210. (Internal citations omitted.)

[119] *Id.* at 215.

[120] *United States v. Alvarez*, 132 S. Ct. 2537, 2552 (2012) (Breyer, J**.,** concurring in the judgment) ("[S]ome such approach is necessary if the First Amendment is to offer proper protection in the many instances in which a statute adversely affects constitutionally protected interests but warrants neither near-automatic condemnation (as "strict scrutiny" implies).")

[121] *273 F.3d 429, 442 (2001).*

intimate privacy.[122] Furthermore, we use technical and legal means to limit the information available to anyone absent a demonstration of proper cause, per the Court's suggestion in *McIntyre*.[123] The question is whether there is a compelling constitutional interest in sexual privacy, a question we turn to next.

B.Sexual Privacy

Non-consensual pornography conflicts with the right to sexual privacy. There is, of course, no explicit right to sexual privacy in the Constitution. Nevertheless, in a series of cases the Court found an implied right.[124]

The most famous such case is *Griswold v. Connecticut*,[125] which invalidated a statute prohibiting the use by or sale of contraceptives to married couples. The Court held that "specific guarantees in the Bill of Rights have penumbras, formed by emanations from those guarantees that help give them life and substance."[126] In particular, it cited the First, Third, Fifth, and especially Fourth Amendments as implying a right to privacy. They further held that the intimate marital relationship was to be strongly protected by this right to privacy: "Would we allow the police to search the sacred precincts of marital bedrooms for telltale signs of the use of contraceptives? The very idea is repulsive to the notions of privacy surrounding the marriage relationship."[127]

The notion of sexual privacy as a fundamental right was extended to all adults, and not just married heterosexuals, in *Lawrence v. Texas*,[128] where the Court struck down Texas 'anti-sodomy statute. They wrote "liberty gives substantial protection to adult persons in deciding how to conduct their private lives in matters pertaining to sex."[129]

---

[122] Arguably, we meet the strict scrutiny requirement as well (*see generally* Stephen A. Siegel, *The Origin of the Compelling State Interest Test and Strict Scrutiny*, 48 AMERICAN JOURNAL OF LEGAL HISTORY 355). However, the need for balancing rights impels us to use exacting scrutiny, which the Court has employed even when dealing with political speech.

[123] *McIntyre v. Ohio Elections Commission*, 514 U.S. 334, 353 (1995) ("We recognize that a State's enforcement interest might justify a more limited identification requirement, but Ohio has shown scant cause for inhibiting the leafletting at issue here.").

[124] These cases, of course, concern government intrusion on sexual privacy, which is not our concern here. Rather, our goal is to show that sexual privacy is a value at the level of a Constitutional right, and hence exerts a considerable balancing weight.

[125] *381 U.S. 479 (1965).*

[126] *Id.* at 484.

[127] *Id.* at 485.

[128] *Lawrence v. Texas*, 539 U.S. 558.

[129] *Id.* at 572.

For that matter, privacy as an abstract concept is viewed as a constitutional right.[130] The Supreme Court recently said as much: "But the *text* of the Fourth Amendment expressly guarantees the 'right of the people to be *secure* in their *persons'*, and our earliest precedents recognized privacy as the 'essence 'of the Amendment—not some penumbral emanation. We have relied on that understanding in construing the meaning of the Amendment."[131]

We are of course not dealing in this article with a government attempt to invade sexual privacy. However, the Court's reasoning in *Griswold*, *Lawrence*, and *Torres* makes it clear that privacy, and especially sexual privacy, is a fundamental right, one that the government should protect. If there is a conflict between it and First Amendment rights, some balance must be struck.[132] Our scheme is intended to protect sexual privacy, with the loss of anonymity if and only if a judge or judges find that some images are indeed non-consensual pornography.

C.Section 230

A simple statute commonly known simply as "Section 230"—more formally, 47 U.S.C. § 230—has been called the statute that "created the modern Internet."[133] Its first operative provision is very simple: "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information

---

[130] There are many different types of privacy. *See, e.g.,* WHO GOES THERE? AUTHENTICATION THROUGH THE LENS OF PRIVACY 63 (Stephen T. Kent & Lynette I. Millett eds., National Academies Press 2003), which describes "bodily integrity", "decisional privacy", "information privacy", and "communications privacy". Our scheme is concerned with bodily integrity, defined in that report as "protect[ing] the individual from intrusive searches and seizures" and information privacy, "protect[ing] the individual's interest in controlling the flow of information about the self to others."

[131] *Torres v. Madrid*, 141 S. Ct. 989, 1002 (2021), citing *Boyd v. United States*, 116 U.S. 616, 630 (1886).

[132] In light of the Supreme Court's holding in *Dobbs v. Jackson Women's Health Org.*, 142 S. Ct. 2228 (2022), this reasoning may need to be revisited in the future. While the Court did not overrule *Griswold* or *Lawrence*, and in fact explicitly denied any intention of doing so ("Finally, the dissent suggests that our decision calls into question *Griswold*, *Eisenstadt*, *Lawrence*, and *Obergefell*… But we have stated unequivocally that '[n]othing in this opinion should be understood to cast doubt on precedents that do not concern abortion.'" *Id.* at 2280), Justice Thomas, in a concurring opinion, explicitly called for that: "For that reason, in future cases, we should reconsider all of this Court's substantive due process precedents, including *Griswold*, *Lawrence*, and *Obergefell*." *Id.* at 2301 (Thomas, J., concurring).

[133] *See generally* JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET (Cornell University Press 2019).

content provider."[134] It allows sites that host content created by others, ranging from YouTube, Twitter, and Facebook to the "Comments" section on obscure blogs, to escape liability if the offending material was "spoken" by these others. Those others might be liable, for anything ranging from libel to obscenity, but the sites themselves are not. At most, they would be obligated to remove it once properly notified of the offense.

This statute has engendered considerable controversy. Indeed, President Trump once vetoed the National Defense Authorization Act in part because it did not amend the statute.[135] Much of the controversy has centered around a different provision: "No provider or user of an interactive computer service shall be held liable on account of … any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected."[136] That is, sites have the right to remove material that they consider "otherwise objectionable," even if it is "constitutionally protected".[137] Trump's "issue with Section 230 came to light this summer after Twitter added warning labels to several of his tweets that alleged mail-in voting is fraudulent,"[138] even though such labeling is likely protected by the First Amendment as well as by Section 230.

---

[134] 47 U.S.C. § 230 (c)(1).

[135] 166 Cong. Recd. H.9150 ("The Act fails even to make any meaningful changes to Section 230 of the Communications Decency Act, despite bipartisan calls for repealing that provision. Section 230 facilitates the spread of foreign disinformation online, which is a serious threat to our national security and election integrity. It must be repealed.").

[136] § 230 (c)(2)(A).

[137] Section 230 was passed in the wake of a court decision in *Stratton Oakmont v. Prodigy Servs. Co.*, 1995 N.Y. Misc. Lexis 229 (Supreme Court of New York 1995). In that case, Prodigy, a then-major online service with multiple "bulletin boards", was accused of permitting allegedly defamatory statements about Stratton Oakmont to appear. The court held Prodigy did exercise editorial functions: "First, PRODIGY held itself out to the public and its members as controlling the content of its computer bulletin boards" *Id.* at *10. This, to the court, made Prodigy more like a newspaper: "with this editorial control comes increased liability" *Id*. at *7, despite Prodigy's claim that 60,000 messages per day made perfect filtering infeasible *Id.* at *8. Section 230 was intended to encourage filtering by eliminating the fear that imperfect filtering would create liability. *See also* KOSSEFF, *supra* note 128, at 60 ("Under the *Stratton Oakmont* rule, these providers had a strong disincentive to take *any* steps to moderate third- party content because any moderation would trigger the responsibility for all third-party content.")

[138] Amanda Macias, *Trump Vetoes Colossal $740 Billion Defense Bill, Breaking with Republican-Led Senate*, CNBC (Dec. 23, 2020), https://www.cnbc.com/2020/12/23/trump-vetoes-740-billion-ndaa-defense-bill.html.

In a controversial move,[139] Congress has already acted once to limit the protections of Section 230 for sites that facilitate prostitution.[140] But what changes might be enacted to satisfy today's political concerns are unclear. Both major parties seem to be concerned about actual moderation decisions but fundamentally disagree on the problem to be solved by amending Section 230: "Democrats have argued that with that protection, companies aren't motivated to remove disinformation. Republicans accuse the companies of using the shield to moderate too much and to take down content that doesn't represent their political viewpoints."[141] Our scheme does not require or even induce companies to engage in any sort of content moderation. Rather, we suggest that they deploy an automated, content-independent origin check.

The more difficult question is why websites should implement our scheme. Let us assume for the moment that the necessary computer code is expressive speech, and hence subject to strict scrutiny. It would almost certainly fail. It is not the least restrictive mechanism for accountability, since it would apply to all websites, and it is not narrowly tailored since it would endanger the anonymity of individuals engaged in political speech, e.g., by posting certain photographs. It does pass the tests under exacting scrutiny, but only if there is another right to be balanced. We address this in two ways. First, we limit the requirement to situations where there is the

---

[139] Mike Masnick, *DOJ Tells Congress SESTA/FOSTA Will Make It MORE DIFFICULT To Catch Traffickers; House Votes For It Anyway*, TECHDIRT (Feb. 27, 2018), https://www.techdirt.com/articles/20180227/15314039324/doj-tells-congress-sesta-fosta-will-make-it-more-difficult-to-catch-traffickers-house-votes-it-anyway.shtml.

[140] Allow States and Victims to Fight Online Sex Trafficking Act of 2017, Pub. L. No. 115-164, § 2(1), 132 Stat. 1253. ("It is the sense of Congress that… section 230 of the Communications Act of 1934 (47 U.S.C. 230; commonly known as the 'Communications Decency Act of 1996') was never intended to provide legal protection to websites that unlawfully promote and facilitate prostitution and websites that facilitate traffickers in advertising the sale of unlawful sex acts with sex trafficking victims."). *See also* Section V.A, *infra*.

[141] David McCabe & Cecilia Kang, *Lawmakers Grill Tech C.E.O.s on Capitol Riot, Getting Few Direct Answers*, NEW YORK TIMES (Mar. 25, 2021), https://www.nytimes.com/2021/03/25/technology/facebook-twitter-google-capitol-riots-hearing.html. *See also* Quinta Jurecic, *The Politics of Section 230 Reform: Learning from FOSTA's Mistakes* (Brookings Institution Mar. 2022). She cites the different political perspectives between the Republican and Democratic parties: "Josh Hawley, then a Republican senator-elect and Missouri's attorney general, suggested in November 2018 that Congress should 'investigate' whether Twitter 'target[ed] political speech'—hinting, wrongly, that an answer in the affirmative would require stripping the platform of its protection from liability" and "Democratic politicians voiced frustration that platforms were leaving too much content *up*—misinformation around the coronavirus, far-right extremism, or lies posted by Trump and his associates."

balancing right of sexual privacy. Second, we provide for a more tailored approach: web sites are free to deploy other mechanisms that provide the requisite degree of accountability, or even to ignore such rules if forfeiting that aspect of Section 230 protections is acceptable to them. Furthermore, we provide for significant judicial scrutiny of requests for deanonymization.

We echo the suggestion made by Professor Danielle Citron and Benjamin Wittes. They propose amending Section 230(c)(1) to read "No provider or user of an interactive computer service *that takes reasonable steps to prevent or address unlawful uses of its services* shall be treated as the publisher or speaker of any information provided by another information content provider *in any action arising out of the publication of content provided by that information content provider."*[142] Professor Citron has gone on to argue that "[t]he determination of what constitutes a reasonable standard of care would consider differences among online entities."[143]

She later modified that proposal to be more specific.[144] Her newer proposal suggests that "reasonable steps" should include a way to report violations or other issues, a process to address those reports, a way to prevent recurring incidents from the same malefactor, minimum logging requirements to enable identification of abusers, and deletion of offending content.[145] Our proposal would address the fourth point: the digital signatures would provide evidence of who first uploaded the offending images, independent of what platform or mechanism they used.[146]

---

[142] *See* Danielle Keats Citron & Wittes, Benjamin, *The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity*, 86 FORDHAM LAW REVIEW 401, 419 (Nov. 2017). [" emphasis in the original.]

[143] Danielle Keats Citron, *Sexual Privacy*, 128 YALE LAW JOURNAL 1870, 1953 (May 2019).

[144] Citron, *supra* note 14, at 39 ("Firms would not know exactly what to do because the proposal did not specify the details of a duty of care.") That is, the proposed statutory text was too vague, leaving it to juries to decide what was or was not reasonable ("Under our proposal, firms wouldn't have been able to predict with certainty whether they were taking the appropriate precautions against different types of illegality. Over time, firms would have learned from experiences with litigation, but they would not have known for sure if their actions fell inside or outside the safe harbor.")

[145] *Id.* at 41.

[146] The privacy-preserving nature of our scheme makes it difficult to use it to block repeated use; an offender could simply get a new subcredential for use with a given site. While previous technical research by one of us has shown ways to blocklist primary credentials (*see* Elli Androulaki et al., *A Real-World Identity Management System with Master Secret Revocation*, No. CUCS-008-10 (Department of Computer Science, Columbia University Apr. 2010)*)*, for complex technical reasons it is not clear if that scheme could be used here. Future technical work is indicated.

D. The Fourth Amendment and "Unreasonable" Searches

Per Part IV, *infra*, we suggest that particularized warrants, supported by probable cause, be required to obtain deanonymization information from the parties to our system. While ordinarily an ordinary *subpoena duces tecum* or grand jury subpoena would suffice, even though the information sought is privacy-sensitive,[147] we believe, following Justice Harlan, that stronger protections are necessary in this situation.[148] Certainly, Congress is free to require such, as it did in the Wiretap Act, where it limited wiretaps to certain specific crimes and where less intrusive investigative techniques are infeasible.[149] We go further and suggest that in this situation, a warrant may be constitutionally mandated. The issue here is not protecting criminal suspects, who apparently have less of a right to anonymity;[150] rather, we wish to protect individuals who may not be credibly suspected of a crime. That is, our scheme provides a mechanism to unmask many different people, regardless of what they may or may not have done; we suggest that strong judicial scrutiny is the best protection against abuse.

The Fourth Amendment[151] is notoriously vague on when warrants are required. It protects people against "unreasonable searches and seizures" without ever defining what "unreasonable" means. We suggest that a search whose purpose is to violate another constitutional right, to wit the right to anonymity, is *a priori* unreasonable unless authorized by a neutral magistrate. That is, the ability to pierce the veil of anonymity around someone should not be at the whim of, say, a law enforcement officer or prosecutor. This is especially true given that, as noted, some images may be

---

[147] *See United States v. Miller*, 425 U.S. 435, 440 (1976), holding that a warrant was not required to obtain access to an individual's bank records ("On their face, the documents subpoenaed here are not respondent's 'private papers.' Unlike the claimant in Boyd, respondent can assert neither ownership nor possession. Instead, these are the business records of the banks.").

[148] *Katz v. United States*, 389 U.S. 347, 361 (1967) (Harlan, J., concurring) ("[T]he invasion of a constitutionally protected area by federal authorities is, as the Court has long held, presumptively unreasonable in the absence of a search warrant.")

[149] 18 U.S.C. § 2516 (3) ("(a) there is probable cause for belief that an individual is committing, has committed, or is about to commit a particular offense enumerated in section 2516 of this chapter; … (c) normal investigative procedures have been tried and have failed or reasonably appear to be unlikely to succeed if tried or to be too dangerous.").

[150] Per Kosseff, "the courts suggest that a criminal suspect or target of a grand jury investigation may be less likely to succeed in preventing a court from allowing the unmasking," *supra* note 67, at 163.

[151] U.S. Const. Amend. IV.

perceived as inducing improper reprisal by law enforcement.[152] Indeed, restraining improper behavior by law enforcement is the core purpose of the warrant requirement in the Fourth Amendment.[153]

There is a further reason for a statutory warrant requirement here. The Fourth Amendment protects "persons, houses, papers, and effects;" it is not obvious that one's anonymity is covered by that.[154] Indeed, similar reasoning was behind the Supreme Court's 1928 holding that wiretaps did not require search warrants,[155] though this holding was later overruled.[156]

Normally, a request for ordinary business records would be via subpoena.[157] We assert, though, that this situation is different. The Court has held that business records that reveal exceptionally sensitive information are more protected. The classic case is *Carpenter v. United States,*[158] where location information is protected. Here, we are contemplating infringing on the First Amendment right to anonymous speech. Specifically, by invoking the Particularity and Probable Cause clauses of the Fourth Amendment, we narrowly tailor the intrusion on First Amendment liberties. We discuss this in more detail in Part V.A, *infra*.

III. TECHNICAL BACKGROUND

A. Overall System Design

We first give a brief description of the parties to the system and the overall flow.

The essence of our proposal is accountability for uploaded images, even if web sites do not attempt to verify users' identities. However, given the sensitivity of personal information, it is essential to protect the privacy of innocent users. In this section, we outline our privacy-preserving upload

---

[152] *See* note 86, *supra*.

[153] This point is made more strongly in Alex Abdo, *Why Rely on the Fourth Amendment To Do the Work of the First?*, 127 YALE L.J. F. 444, 455 (2017) ("where the First Amendment applies, it would require the government to demonstrate a heightened interest to justify its surveillance").

[154] *But see* the discussion of the First Amendment and anonymity, Part II.A, *supra*.

[155] *Olmstead v. United States*, 277 U.S. 438 (1928) *overruled by Katz v. United States*, 389 U.S. 347 (1967).

[156] *Katz*, 389 U.S. 347.

[157] *United States v. Miller*, 425 U.S. 435, 440 (1976). While image signatures are relatively meaningless to web sites, the information necessary to unmask them is contained in the business records of the certificate authority, the deanonymization agent, and the identity provider, per Part III, *infra*.

[158] *Carpenter v. United States*, 138 S. Ct. 2206, 521 (2018) ("We decline to extend *Smith* and *Miller* to cover these novel circumstances. Given the unique nature of cell phone location records, the fact that the information is held by a third party does not by itself overcome the user's claim to Fourth Amendment protection.").

scheme to create accountability while preserving privacy. We first describe the necessary technical underpinnings of our scheme—cryptographic certificates and digital signatures, Exif metadata in digital photographs, and an advanced cryptographic protocol designed by researchers Jan Camenisch and Professor Anna Lysyanskaya. We then explain how these pieces fit together in our image uploading flow, which encrypts and preserves an uploader's credentials in metadata, and our uncovering flow, which decrypts credentials.

Users—that is, anyone who wishes to upload images to a participating web site—must first register with an *Identity Provider*. To do this, they must provide evidence of their identity similar to what would be needed to notarize documents.[159] This step is partially automated and partially reliant on manual action by the users.

A *participating web site* is one that implements the technical mechanisms we describe. Their incentive to participate is to gain the full protections of Section 230 for user-generated content, as described by Prof. Danielle Citron and Benjamin Wittes.[160] (There may be other suitable mechanisms for gaining such protection per their analysis; we discuss only ours.) Sites that do not accept image uploads, or ones that employ their own filtering mechanisms, *e.g.*, a news site, need not participate.

To upload an image to participating web sites, the user's browser must first digitally sign the image. This requires the browser to first engage in a dialog with the Identity Provider and a *Certificate Authority*. The Certificate Authority issues a privacy-preserving *certificate* to the browser. The dialog is fully automated and no user interaction is needed; it is fully transparent to the user, save for a short delay the first time an image is uploaded to a given site. The web site validates that the signature is correct for the uploaded image.

If an image that is believed to be non-consensual pornography is referred to law enforcement, an officer (by way of a *Law Enforcement Portal)* obtains certain encrypted information from the Certificate Authority and sends it to a *Deanonymization Agent*. The Deanonymization Agent decrypts this information, yielding a user's pseudonym,[161] and returns it to the Portal. Finally, the Portal sends the pseudonym to the Identity Provider. It in turn provides the user's identity, per the original registration. Suitable legal process is necessary for this deanonymization to take place.

---

[159] *See* Part V.C, *infra*, for a discussion of the social and economic issues surrounding registration.

[160] Citron & Wittes, Benjamin, *supra* note 137.

[161] This pseudonym is a large random number, not a screen name or the like, and need not be shown to, let alone memorized by, the user.

To summarize:

1. A user must first obtain a privacy-preserving credential from an Identity Provider (IDP); the IDP knows who they really are but does not include that in the credential. Only the IDP knows the user's real identity.
2. These credentials are transparently used to obtain cryptographic certificates from a Certificate Authority (CA) for each image-accepting website that the user wishes to visit.
3. The CA also retains an encrypted file that can contribute to identifying the user.
4. The certificate is used to digitally sign images being uploaded to websites that participate in this scheme.[162] If a judge finds, in an *ex parte* proceeding, that an image is, in fact, nonconsensual pornography, a warrant can be issued to revoke the anonymity of the person who uploaded it.
5. The CA supplies that encrypted file, the Deanonymization Agent decrypts it, and the IDP supplies the uploader's real identity. At that point, normal law enforcement actions can commence.

Revoking anonymity thus requires the cooperation of three independent parties; this protects privacy from ordinary commercial actors and impedes improper behavior by law enforcement.

E.Certificates

Cryptology is an ancient discipline. Some authors believe it goes back 4,000 years.[163] While there have been many different ways that encryption has been implemented, a cardinal modern principle is that performing encryption requires both an algorithm and a "key."[164] It has been recognized since 1883 that the secrecy of the key should be the sole determinant of a system's strength; the algorithm itself may be known to the enemy.[165]

---

[162] These terms are explained in more detail in Part III, *infra*.

[163] *See generally* DAVID KAHN, THE CODEBREAKERS, ch. 2 (Scribner 2nd ed. 1996).

[164] Keys are discussed in all cryptography textbooks; *see, e.g.,* JONATHAN KATZ & YEHUDA LINDELL, INTRODUCTION TO MODERN CRYPTOGRAPHY § 1.2 (CRC Press Third ed. 2021) ("Security of all classical encryption schemes relies on a secret—a *key*—shared by the communicating parties in advance and unknown to the eavesdropper.")

[165] Auguste Kerckhoffs, *La Cryptographie Militaire*, 9 JOURNAL DES SCIENCES MILITAIRIES 5, 12 (Jan. 1883) ("Il faut qu'il n'exige pas le secret, et qu'il puisse sans inconvénient tomber entre les mains de l'ennemi"—roughly, "it [the cryptosystem] must not require secrecy, and must be able to fall into enemy hands without inconvenience.")

Traditional cryptographic mechanisms required that the sender and the recipient of an encrypted message share the same key: if the sender were to encrypt a message with a key, the recipient would not be able to read it unless they knew the same key. The need to securely share keys was a considerable restriction, but no alternative was publicly known until 1976.[166] Whitfield Diffie, then a graduate student, and Professor Martin Hellman conceived of what is now known as "public key cryptography." Their essential notion was to posit a system where there were two keys, a "public" key (which anyone could know) that is used to encrypt messages, and a closely held "private" key to decrypt them. Furthermore, the private key can be used to create a "digital signature," a statement of the origin of a document which can only be created by someone who knows this closely held key. The signature can, though, be verified by anyone who know the public key—and that can be widely distributed.

Most uses of public key cryptography rely on public key certificates, electronic documents issued by a trusted *Certificate Authority (CA)*; they are used to prove ownership of public keys.[167] A certificate is a statement of a name and that entity's public key, all digitally signed with the CA's private key. There can be other data in a certificate, including an expiration date: it cannot be used after that period.[168] Anyone can verify that signature on the certificate if they know the CA's public key. Certificates today, including all common forms of encryption on the World Wide Web, generally use a format known as X.509. [169] When a user connects to, say, https://www.supremecourt.gov, the web site returns the certificate for www.supremecourt.gov; the browser checks the certificate using a list of

[166] Diffie & Hellman, *supra* note 29. The concept was actually invented a few years earlier at the Government Communications Headquarters (GCHQ), the British equivalent of the NSA; *see* James Ellis, *The Possibility of Secure Non-Secret Encryption* (GCHQ Dec. 1969) and Clifford Cocks, *A Note on "Non-Secret" Encryption'* (GCHQ Nov. 1973).

[167] Certificates were invented by an undergraduate as part of his senior thesis (Loren M. Kohnfelder, *Toward a Practical Public-Key Cryptosystem* (Department of Electrical Engineering, Massachusetts Institute of Technology May 1978).). A public key certificate is an electronic document issued by a trusted third-party Certificate Authority used to prove ownership of a public key. *See* Andreas Pfitzmann et al., *A Terminology for Talking about Privacy by Data Minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management* 98, 28.

[168] A variety of technical information, such as the precise algorithms used, is also included in certificates.

[169] *Information Technology — Open Systems Interconnection — The Directory: Public-Key and Attribute Certificate Frameworks*, No. X.509 (International Telecommunications Union Oct. 2019).

trusted CAs built into the browser or the user's operating system.[170]

Our scheme relies on special-purpose CAs. Rather than tying a key to a name, these CAs associate a public key with a subcredential issued by the identity provider. Uploaded images include both a digital signature of the image and the accompanying certificate. The signature and certificate will be stored in a clearly defined data field in the image's metadata. Web sites will know to not wipe this field because it contains information that can be be used to trace an image back to the uploader. To be sure, malicious or non-conforming web sites may delete this information. There is little that can be done about malicious sites, though they may as a consequence lose Section 230 protection. For sites that are benignly non-conforming, it will simply take time for them to adopt this extra step.

B.Exif Data

Exchangeable Image File Format (Exif), which defines a file standard for metadata tags used by digital cameras, was originally developed in 1995 to promote consistency between photographic devices.[171] Digital cameras record identifying details in Exif data such as camera make and model, date and time information, location information, and even the owner's name.[172] In this way, it is the modern counterpart to writing the dates or locations on the back of photographs in order to remember when, where, and how they were taken—but just like this information, it can be faked by knowledgeable users.

Many users are unaware of how much tracking information can be contained in a single photo. Although Exif data may be replaced through image editors, many users do not think about editing image metadata to protect their privacy before uploading their images online. To protect against this, social media sites usually scrub the metadata from photos to protect

---

[170] If the CA were not trustworthy, it could issue fraudulent certificates for web sites, thus defeating the intent of the certificate system. Such problems have happened on the web when CAs were hacked; *see, e.g., Black Tulip: Report of the Investigation into the DigiNotar Certificate Authority Breach*, No. PR-110202 (Fox IT Aug. 2012), *available at* http://www.rijksoverheid.nl/bestanden/documenten-en-publicaties/rapporten/2012/08/13/black-tulip-update/black-tulip-update.pdf.

[171] Version 1.0. Published October 1995. https://www.loc.gov/preservation/digital/formats/fdd/fdd000146.shtml.

[172] Official Exif 2.32 tag specifications. https://web.archive.org/web/20190624045241if_/http://www.cipa.jp:80/std/documents/e/DC-008-Translation-2019-E.pdf. Owner name is not set by default, but can be set by photo editing software.

users' information.[173] This is a common practice in the privacy community to prevent users from inadvertently leaking identifying information.[174] However, since Exif data has also demonstrated effectiveness in helping law enforcement track NCP, uploader accountability may be lost as a consequence of wiping all Exif data on upload.[175]

Exif processing and manipulation must be handled correctly by web servers in order to protect user privacy. Social media apps have supplemented the increased accessibility of powerful cameras on mobile devices in the last decade by making it more convenient than ever for users to upload straight from their mobile devices in seconds.[176] Software is available to view and edit Exif fields and add user-customized fields to media files.[177] The extensibility of Exif data-fields—it is possible to add

---

[173] Apparently as a response to stories warning people against inadvertently sharing location information when posting photos online, social media platforms Twitter and Facebook have publicized that they do not make Exif data available on the uploaded photo and that users can decide to add Facebook or Twitter tags if they wish to add location. *Facebook Security - Posts | Facebook*, https://www.facebook.com/security/posts/sharing-photosthe-facebook-security-team-has-received-a-number-of-questions-abou/10151511111506886/ (last visited Apr. 2, 2021); *How to Tweet Pictures or GIFs | Twitter Help* § What happens to the Exif data for my photo?, https://help.twitter.com/en/using-twitter/tweeting-gifs-and-pictures (last visited Apr. 2, 2021); Flickr, a photo-sharing platform aimed at photography enthusiasts, informs users of what information Exif data can expose, but does not actively remove it—photographers often want to know what settings were used for a particular picture. *Exif Data FAQ*, https://help.flickr.com/en_us/understand-flickr-exif-data-r1ge02Xo1X (last visited Aug. 10, 2021) ("Showing your Exif data will allow those admiring your photos to learn about the camera, lens, aperture, and other settings that were used.").

[174] A possible reason for web sites to completely delete Exif data is because web sites do not dictate the standards for how Exif should be used, so it is safer to remove data from all fields than to risk leaking data when field names are updated. Additionally, Exif data fields often contain redundant information. For example, on photographs exported by Adobe's Lightroom Classic Version 12.0.1 on MacOS, the interval during which an image was captured is given three times, as "Exposure Time", "Shutter Speed Value", and "Shutter Speed." Similarly, location appears as both "GPS Position" and the pair "GPS Latitude" and "GPS Longitude". The photographer's name can appear at least six times: "Artist", "Copyright", "By-line", "Copyright notice", "Creator", and "Rights". (Experiments by one author.) Therefore, in order for web sites to preserve anonymity, it is best practice to delete all metadata.

[175] Salvatore Paladino, *Exif: A Format Is Worth A Thousand Words*, TECHBEAT (Winter 2007), https://www.ojp.gov/pdffiles1/nij/nlectc/218283.pdf.

[176] *How Do I Share Photos on Facebook? | Facebook Help Center*, https://www.facebook.com/help/mobile-touch/174641285926169 (last visited Aug. 10, 2021); *How to Tweet Pictures or GIFs | Twitter Help*, *supra* note 168, § To post a photo or GIF in a Tweet.

[177] *Id.*

new fields to the standard or to images—allows us to use these fields as indicators to web sites,[178] telling them what to refrain from wiping.

Our scheme involves adding a chain of digital signatures to the Exif field in uploaded images. The extensibility and customizability of Exif fields provide an opportunity for web sites to identify what information is important enough to persist when metadata is processed. The data field will be uniquely named so a web site can identify that it contains information that should not be wiped during upload.[179]

C. Camenisch-Lysyanskaya Credentials

Our primary tool for protecting privacy is a cryptographic protocol designed to preserve user-anonymity but with an ability to deanonymize users if further steps are followed. Our design requires multiple authorities to cooperate before user identity is revealed.[180] Not only is user privacy protected, but accountability is preserved. We base our design on a credential system with revocable anonymity outlined in a paper by Jan Camenisch and Professor Anna Lysyanskaya (hereinafter Camenisch-Lysyanskaya or CL Credentials).[181]

CL Credentials have privacy preserving properties that allow user identities to remain anonymous during use. Users can ultimately lose their anonymity if an unlawful image is discovered to be linked to their credential on upload, albeit with the cooperation of multiple parties. These credentials are provably unlinkable in normal situations to the identities of users but can be verified as belonging to the underlying identity that originally requested them through complex zero-knowledge cryptographic proofs demonstrated in the original Camenisch-Lysyanskaya paper.[182]

For the purposes of this paper, we will not discuss the mathematics behind these credentials. For the most part, we can treat the credentials, consisting of a Pseudonym, Primary Credentials, Subcredentials,

---

[178] *Id.*

[179] Software like Exiftool can then be used to view, edit, or selectively wipe data fields. *ExifTool by Phil Harvey*, https://exiftool.org/ (last visited Mar. 17, 2021).

[180] Full details are given in a companion technical paper, PJacob Gorman et al., *Privacy-Preserving Accountability for Non-Consensual Pornography*, Draft (2023).

[181] Camenisch & Lysyanskaya, *supra* note 28.

[182] *Id.* The concept of zero-knowledge proofs was introduced by S Goldwasser et al., *The Knowledge Complexity of Interactive Proof-Systems*, Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing 291 (Association for Computing Machinery 1985).

Deanonymization String,[183] and Certificate, as opaque objects[184] with a few important external properties:

1. The Pseudonym and Primary Credential are associated with the User's identity, but only the Identity Provider knows this association.[185]
2. The Sub-credential is verifiably derived from a valid Primary Credential.[186]
3. The Deanonymization String, if decrypted, can be used to obtain the Pseudonym.[187]
4. The Certificate is associated with a public key and encrypted pseudonym.[188]
5. The Deanonymization Agent can unencrypt the Deanonymization String.[189]

---

[183] For clarity, we use the terms Primary Credential and Subcredential, though they are not used in the original Camenisch-Lysanskaya paper. We use the phrase Deannoymization String where the paper uses Revocation String; that latter phrase is, in our experience, too easily confused with more common uses of "revocation" when discussing cryptography.

[184] By "opaque object", we mean that the details of what they contain and how they work are not important to the legal discussion. Rather, what is important is their external properties.

[185] The User and IDP work together to generate a Pseudonym. The Pseudonym is used to create a Primary Credential. The Primary Credential can be used to generate Subcredentials that are not linkable to the Primary Credential or each other. Although they are privacy preserving, they are provably valid and retain the ability to authenticate a user's identity. The technical details are exceedingly complex, so we will not try to explain it here. For a thorough explanation, refer to the Camenisch-Lysyanskaya paper. Camenisch & Lysyanskaya, *supra* note 28.

[186] The Subcredential can be presented to a third party (the CA) to verify they are valid without being able to see whose they are, and which Primary Credential generated them. This is done using zero-knowledge proofs, a way to prove something without giving away what it is. The effect is similar to being presented a notarized identifying document and reading it using a cardboard with a peephole so that just the notary's seal is shown. We can verify the authenticity of the identity, without knowing the contents. The CA knows it is valid but does not know who created it. The Subcredential also creates the Deanonymization String, which is a separate string that contains information about the (encrypted) Pseudonym that was used to create the original Primary Credential.

[187] The Deanonymization String can be presented to a third party (the CA) to verify it is derived from the Subcredential without being able to see whose they are. Again, zero knowledge proofs are used to verify the validity of the string.

[188] The CA creates a Certificate which is an association between the user and their public key. Instead of the user's identity, the public key will be associated with a constant protocol-identifying string. The CA will store the Deanonymization String with the Certificate's serial number to link the Pseudonym with the Certificate.

[189] Law Enforcement may ask the Deanonymization Agent to decrypt the Deanonymization String to reveal the Pseudonym.

6. The Pseudonym is linked to the person's real identity.[190]

F.Parties to the System

In this section, we will introduce the five main parties: Users and Browsers, Identity Provider, Certificate Authority, web sites, Deanonymization Agent. We will discuss the assumptions made about each party. A sixth component, the Law Enforcement Portal, is added for convenience but is not an essential piece.

The Identity Provider verifies identities with approved government documents, admissible to at least the standards required by a notary public.[191]

Users (and their Browsers) are ordinary individuals who mainly will not be uploading NCP. (Users may try to fake their identities by providing a fake document to the Identity Provider; this would be handled by ordinary law enforcement means.) Users may also try to circumvent the browser and upload without a signature or try to spoof their signature with an invalid certificate of their authenticity. We assume that users will do all they can to breach normal operation: if all users were completely law-abiding, there would be no need for our scheme. Implementations of our scheme, then, must detect and reject such uploads. Specifically, Certificate Authorities must verify that credentials are valid and web sites must verify that certificates in images are valid and that web site digital signatures are correct.

The Certificate Authority accepts and verifies the Sub-credential's validity and issues an industry-standard X.509 Certificate associated with a cryptographic key. CAs must be trusted by web sites,[192] since the

---

[190] Law Enforcement may ask the IDP to decrypt the Pseudonym to reveal the User's identity.

[191] The precise standards that a notary public must follow are a matter of state law. New York's standards are given in *Satisfactory evidence of identity*, 19 NYCRR § 182.5. Those standards include the usual government-issued photo ID, but also two other documents with the person's name, address, and signature, personal knowledge of the person by the notary, attestation by suitable witnesses, and more. California has similar requirements (Cal. Civ. Code § 1185(b)), but the state also requires that for some notarizations, a log of the signer's thumbprint be kept, Cal. Gov. Code § 8206(a)(2)(G) ("If the document to be notarized is a deed, quitclaim deed, deed of trust, or other document affecting real property, or a power of attorney document, the notary public shall require the party signing the document to place his or her right thumbprint in the journal.")

[192] Since a certificate authority is vouching for an identity, it must by definition be trusted by the parties that rely on it; *see* note 159, *supra*. For an explanation of trust models for certificate authorities, *see, e.g.,* R. Perlman, *An Overview of PKI Trust Models*, 13 IEEE

Certificates they issue are used to digitally sign user-provided images before they are uploaded.[193]

Web sites act as service providers and do not act as content providers, meaning they do not contribute materially to the content they host.[194] Such web sites cannot be treated as publishers of content and enjoy immunity under Section 230. However, if they choose to participate in this protocol, they must verify the signatures in all uploaded images or they could be fooled by missing, spoofed, or invalid signatures.

The Deanonymization Agent is a trusted neutral party that is capable of decrypting Deanonymization Strings, though in our scheme it should only do so when presented with a suitable legal process. It is not involved in day-to-day operation until there is some legal intervention.

The IDP, CA, and web sites should be expected to have reliable response in real time. The Deanonymization agent does not need to respond in real time; its response time may be gated by the necessity for legal process.

The Identity Provider, Certificate Authority, and Deanonymization Agent are all regulated and possibly licensed parties because they must be trusted to carry out their parts in the operation without collusion.[195] To minimize collusion, they should also not be part of the same corporation. These parties can be described as "honest but curious:" they are trusted to carry out their roles but they will look at anything that's unencrypted.[196] This lack of collusion is important for preserving the 3-party operation to deanonymize someone. These parties are assumed to keep adequate records and cooperate with the legal process.

The Law Enforcement Portal is a simple web site that police will use to investigate cases of NCP. It will engage in the low-level dialogs with the active entities, the CA, the Deanonymization Agent, and the Identity Provider, to learn who uploaded images.

NETWORK 38, 38 (1999) ("I f Bob trusts a particular CA and knows that CA's public key, he can securely know Alice's public key if he can obtain a certificate signed by that CA certifying Alice's public key as belonging to the name Alice.")

[193] Due to the nature of Subcredentials, they cannot be used directly to sign images. *See* Camenisch & Lysyanskaya, *supra* note 28.

[194] They are thus not "speakers" according to Section 230 (47 U.S.C. § 230 (c)(1).).

[195] *See* Part IV, *infra,* for a more complete discussion of these requirements.

[196] "Honest but curious" is a standard term in the cryptographic literature; *see, e.g.* 2 ODED GOLDREICH, FOUNDATIONS OF CRYPTOGRAPHY 603 (Cambridge Univ Press 2001).

D.Normal Operation

Before anything, the user must use a browser that implements this scheme. This is important for the next step, during which the user verifies their identity with an IDP. The user may then visit a web site where they can upload images. If they attempt to visit the web site before their pseudonym and primary credential are created by their browser and IDP, they would be redirected to the IDP to complete their identity verification. This step is a one-time operation, and would be required at most every few years.[197]

Identity verification must authenticate the user to a standard similar to that used by notaries. Part of the duties of a notary certified by a state government is to screen signers of important documents for their true identities. To protect against fraud, identity verification typically requires physical presence at a notary, which the IDP would require as well. Identities verified at an IDP would be admissible to at least the standard of a notary, in order to provide enough credible evidence to inculpate someone.[198] The identity provider stores a list of pseudonyms—large, random numbers—and real identities.

For technical reasons, images cannot be signed directly with CL credentials. Accordingly, in order to upload an image to a cooperating web site,[199] the browser must obtain a sub-credential from the IDP and present it to a CA. This happens automatically, without user interaction. The browser also works with the IDP to generate a Deanonymization String that contains a "blinded" Pseudonym.[200] The Sub-credential and the Deanonymization String are validated by the Certificate Authority (CA). The CA will then generate a certificate with an arbitrary or random username not linked to anything, but with the certificate's unique serial number stored in a database with the Deanonymization String. Once the certificate is stored in the browser, the user will not experience delays on subsequent visits until their certificate expires and they need to request a new one. The browser will

---

[197] It is normal practice for credentials such as driver's licenses to expire after a few years. A similar requirement would be reasonable for primary credentials. Note, though, that the expiration period is a matter of years, not weeks or months.

[198] A detailed rationale for this requirement is given in Part IV, *infra*.

[199] If a web site wishes to participate in this scheme—and as discussed, this is not mandatory—pages from which images can be uploaded contain a signal that is recognized by the user's browser.

[200] In cryptography, a blinded object is one that is not comprehensible to the viewer but still has all necessary properties. The concept was originally introduced by David Chaum (*Blind Signatures for Untraceable Payments*, Advances in Cryptology 199 (David Chaum et al. eds., Springer US 1983)) where he described "blind signatures": the ability to sign something without knowing what it is you are signing.

include a digital signature of the image and the associated Certificate in each upload's Exif data as a "SignatureBlock" as proof of the user's identity. Although users are not assumed to behave honestly or correctly, certificates from a trusted CA are reliable and can be verified as genuine by the web site.

The web site will verify the upload's SignatureBlock before accepting it. If there are any inconsistencies with the signature or the certificate, the web site will reject the upload and notify the user of the reason it was rejected.

If the image is accepted by the server and law enforcement detects that it is NCP, all of the information needed to trace the image back to the original uploader (as well as subsequent uploaders 'identities) is in place. This process is shown in Figure 1.

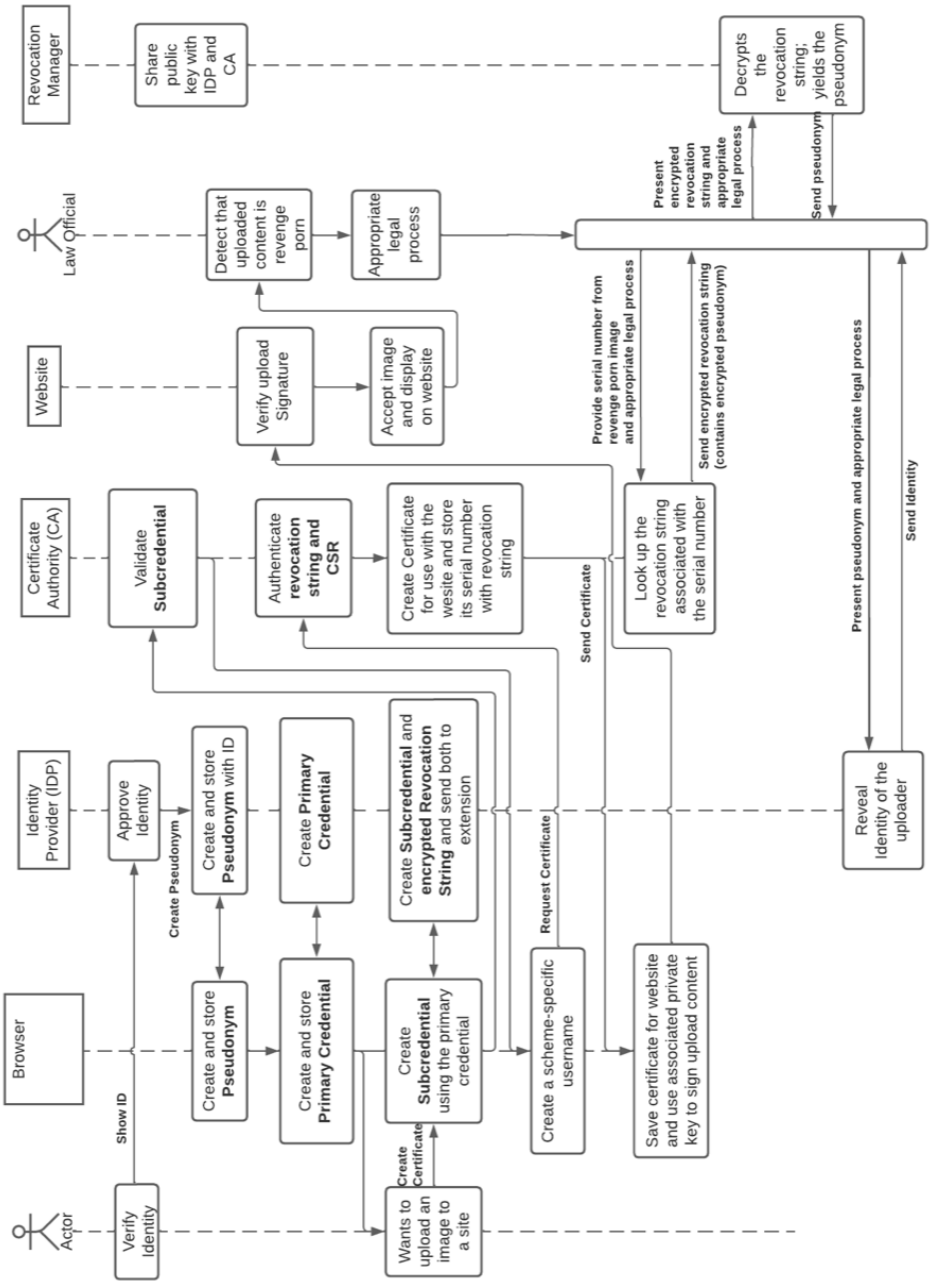

Figure 1: Overall data flow

E.Law Enforcement Actions

Under this system, law enforcement can uncover identities after following the proper legal process when they discover that an upload is revenge porn. First, they retrieve the certificate embedded in the SignatureBlocks field of the Exif data. Then law enforcement must take the serial number of the certificate and after going through the appropriate legal process, present it to the CA to request the deanonymization string used when the certificate was requested. Law enforcement will present the deanonymization string to the Deanonymization Agent, again with appropriate legal process. The Agent will cooperate by returning the pseudonym, which was created at the IDP. Then law enforcement will present the pseudonym to the IDP (with the appropriate legal process) and the IDP will reveal the identity of the uploader.

Since subsequent signatures on an upload are chained together in the SignatureBlocks Exif field, law enforcement can follow this protocol for every signature in the chain. This allows the possibility of judicial action to be taken against every re-uploader of that image, if appropriate, specifically including the original uploader.

Because the steps involved in ascertaining identities are rather involved, we suggest the creation of an application we call the Law Enforcement (LE) Portal. The LE Portal is not a privileged component, in that no one has to trust it; it can be implemented as a local application in a police station, one or more web sites, etc. It simply automates the necessary technical queries.

The portal has to perform several different steps, with arbitrary delays between steps. First, it must accept uploads of arbitrarily many images and extract the SignatureBlocks from them; this yields a set of certificates for each such image. Possession of an image is presumed to be sufficient authority for the extraction; like any other image metadata, it is embedded in the file and could easily be extracted without recourse to an official Portal. There is thus no manual processing, and a response would likely be immediate.

The next step, though, is more complex: asking a CA to supply the deanonymization string for each certificate. As discussed in Parts IV and V.A, *infra*, such requests require judicial approval. Thus, the Portal must send a copy of the court order to the CA. Furthermore, since we posit that the CA should have the right to oppose such an order, the images in question must be sent to the CA as well. This step could therefore be subject to a long delay while the images are analyzed by the CA and perhaps while litigation takes place.

The third and fourth steps have similar characteristics: the Deanonymization Agent must be sent the deanonymization strings, the

court order, and the images before it responds with a pseudonym. The IDP would receive the pseudonym and the requisite documents before supplying the user's identity. Again, there can be an indefinite delay before there is a response. The LE Portal must be implemented in such a way as to account for all of these delays.

To prevent nuisance submissions to the CA, etc., it is desirable that there be some sort of authentication of requests. There are several ways to approach this. One is to piggyback on existing law enforcement authentication systems. The National Crime Information Center, a set of databases run by the FBI, "is a collaboration between the FBI and federal, state, local, and tribal criminal justice users" and employs authentication mechanisms to validate users.[201] Access for civil suits would presumably be provided by the courts after a suit was filed.[202]

Especially in the case of civil suits, it is desirable that court orders be authenticated. This is technically feasible; it is also desirable, given the incidence of forged court orders and emergency surveillance requests.[203] A bill was introduced in Congress in 2021 to combat the problem, but it did not pass.[204]

IV. Operational Analysis

Ultimately, our scheme is a way to revoke someone's anonymity and hence invade their privacy.[205] Here, however, we are advocating it as a way to deter or punish a more serious invasion of privacy.[206] We therefore rely on legal processes to provide suitable protection for all parties.

---

[201] *National Crime Information Center*, https://le.fbi.gov/informational-tools/ncic (last visited July 14, 2023). Such access can, of course, be abused, *see, e.g., United States v. Valle*, 807 F.3d 508 (2nd Cir. 2015). According to J. Van Duyn, Automated Crime Information Systems 18 (TAB Books 1991), "[b]efore a terminal operator can make an inquiry, entry, modification, or cancellation, into any NCIC database, he must supply the system a NCIC-assigned ORI [Originating Agency Identifier], identifying the originating agency. He must also give his authorization code."

[202] *See* note 88, *supra*.

[203] Volokh, *Shenanigans (Internet Takedown Edition)*, 2021 U.L.R. 237 (2921); Brian Krebs, *Fake Emergency Search Warrants Draw Scrutiny from Capitol Hill*, Krebs on Security, https://krebsonsecurity.com/2022/03/fake-emergency-search-warrants-draw-scrutiny-from-capitol-hill/ (last visited Jul. 14, 2023).

[204] Digital Authenticity for Court Orders Act of 2021, S.2547, 117th Congress Sess. (2021).

[205] *See, e.g.,* Kosseff, *supra* note 67, at 176 ("But as seen in the Malibu Media cases and so many other attempts to unmask anonymous Internet users, anonymity is more than just about the ability to express oneself. It is a privacy concern.").

[206] Daniel J. Solove, The Future of Reputation 141 (Yale University Press 2007)

As noted, what is necessary to deanonymize an image is record retrieval and/or computation from three different parties. We must ensure that the necessary disclosures and actions occur if and only if appropriate. That is, we must ensure that these parties do cooperate with proper legal requests; conversely, we must ensure that requests from law enforcement are indeed proper. We address the second issue first.

If there were no other protections required, the legal flow would be simple in criminal cases: law enforcement would request that the deanonymization agent perform the necessary computations, perhaps with a back-up order under the All Writs Act.[207] The certificate authority and the Identity Provider need only do a database lookup; a simple subpoena would suffice. Professor Kosseff's observation about online anonymity in criminal cases notwithstanding,[208] we regard this as insufficiently privacy-protective. If nothing else, there are often public perceptions and perhaps the reality of police harassment[209] that must be guarded against. Furthermore, there may be civil suits where judicial oversight is crucial, and a mere subpoena is inadequate.

It is instructive to look at some different levels of oversight for different types of requests. In national security investigations, the FBI can request certain records with no judicial oversight at all.[210] However, the target of such orders can request judicial review.[211] In an ordinary criminal investigation, the same sort of material is only available after showing a judge specific grounds for the order's issuance.[212] If a pen register or a trap-

("Thus anonymity is a form of privacy protection, yet it can also facilitate privacy violations.").

[207] 28 U.S.C. § 1651.

[208] *See* KOSSEFF, *supra* note 67, at 158.

[209] *See, e.g.,* Chelsea Torres & Nicholas Lopez, *Mother Speaks about Saving Her Children from Uvalde Gunman*, FOX 29 (Jun. 26, 2022), https://foxsanantonio.com/newsletter-daily/mother-speaks-about-saving-her-children-from-uvalde-gunman ("Ever since that harrowing day, Gomez says she has faced scrutiny from law enforcement, even at her own home, 'the other night we were exercising and we had a cop parked at the corner like, flickering us with his headlights.'")

[210] 18 U.S.C. § 2709 (a) ("A wire or electronic communication service provider shall comply with a request for subscriber information and toll billing records information, or electronic communication transactional records in its custody or possession made by the Director of the Federal Bureau of Investigation").

[211] *Id.* (d).

[212] 18 U.S.C. § 2703 (d) ("A court order for disclosure under subsection [b] or [c] may be issued by any court that is a court of competent jurisdiction and shall issue only if the governmental entity offers specific and articulable facts showing that there are reasonable grounds to believe that the contents of a wire or electronic communication, or the records or other information sought, are relevant and material to an ongoing criminal investigation.").

and-trace device is used,[213] the investigating officer merely has to certify to the judge that there are adequate grounds; the judge does not get to rule on their adequacy.[214] A search warrant, of course, requires the full protection of the Fourth Amendment: "no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized."[215] But stronger requirements are possible; the Wiretap Act specifies that taps are only permissible for certain specified crimes and under certain circumstances, and that is in addition to the Fourth Amendment requirements.[216] What standard should be used for deanonymization?

Courts have been dealing with Internet deanonymization issues for years, often in defamation cases. There have been no U.S. Supreme Court cases that squarely address the issue; instead, for guidance we turn to two influential state cases, *Dendrite Int'l, Inc. v. Doe No. 3*[217] and *Cahill v. Doe*.[218] Both cases contain frameworks for deciding on whether to deanonymize an Internet user and under what conditions.

Dendrite, a case decided by an intermediate New Jersey appellate court, involved a software company that felt it was defamed by pseudonymous posters on a Yahoo! Finance bulletin board, possibly to drive down the stock price. The company also alleged theft of trade secrets, based on the claim by some posters that they had access to non-public information about the Company. The court set out a four-pronged process for deciding if a purported defamer should be identified:[219] (1) private notification to the subject, so that they can, if they wish, defend their anonymity; (2) identification of the exact statements that are allegedly actionable; (3) review by the court to ensure that the complaint would survive a motion to dismiss, i.e., that there is a *prima facie* cause of action; and (4) that the

---

[213] A pen register records what numbers a subscriber dials; a trap-and-trace device records calling numbers. *See* 18 U.S.C. § 3127 (3) and § 3127 (4).

[214] 18 U.S.C. § 3122 (2) ("include… a certification by the applicant that the information likely to be obtained is relevant to an ongoing criminal investigation being conducted by that agency.")

[215] U.S. Const. Amend. IV.

[216] 18 U.S.C. § 2516 (3) ("(a) there is probable cause for belief that an individual is committing, has committed, or is about to commit a particular offense enumerated in section 2516 of this chapter; … (c) normal investigative procedures have been tried and have failed or reasonably appear to be unlikely to succeed if tried or to be too dangerous."). For the Fourth Amendment requirements for wiretaps, *see Katz v. United States*, 389 U.S. 347 (1967).

[217] 775 A.2d 756 (N.J. App. Div. 2001).

[218] 884 A.2d 451 (Del.).

[219] *Dendrite*, 775 A.2d at 760–761.

judge balance the defendant's First Amendment right to anonymity against the strength of the *prima facie* case and whether the case could proceed without such identification.

Cahill was also a defamation case that additionally alleged of invasion of privacy. The underlying facts are complex.[220] Broadly speaking, they involved a dispute between town councilman Patrick Cahill and his next-door neighbor Mayor Mark Schaeffer, the apparent owner of what was either a security camera or a surveillance camera. The dispute soon spilled online. The *Cahill* court considered the *Dendrite* process but decided to simplify it. In particular, it dropped the second and fourth elements.[221] The second element, it felt, was subsumed by the third: a determination that there was a *prima facie* cause of action necessarily included understanding exactly what was defamatory. The court also felt that an explicit balancing was unnecessary, in that was also implicit in the summary judgment standard it adopted for the third element.[222]

We cannot apply either of these processes literally. If nothing else, in our scheme only the identity provider actually knows the identity of the purportedly offending party. Beyond that, both the *Dendrite* and *Cahill* courts suggest that a notice be posted on the message board where the offending statement originally appeared.[223] We cannot assume any such continuity of access. For one thing, an offender may upload pictures to a site and never return. For another, pictures are often downloaded from one site by a third party and uploaded to another site; again, the original offender would not see a notice posted there. Notification is a good idea and the identity provider should be required to do it, but we cannot rely on it for substantive protection at the earlier stages of identification. Rather, the certificate authority and the deanonymization agent could act as the putative offender's proxy in opposing such a motion. There are, however, complex economic issues implicated here; we defer discussion of those to Part V.C, *infra*.

Whether we adopt the *Dendrite* test or the simpler *Cahill* test, we must answer two questions: what is the standard by which purported

---

[220] KOSSEFF, *supra* note 67, at 126–131.

[221] *Cahill*, 884 A.2d at 461.

[222] *Id.*

[223] *Dendrite*, 775 A.2d at 760 ("These notification efforts should include posting a message of notification of the identity discovery request to the anonymous user on the ISP's pertinent message board."); *Cahill*, 884 A.2d at 461 ("In the internet context, the plaintiff's efforts should include posting a message of notification of the discovery request to the anonymous defendant on the same message board as the original allegedly defamatory posting.")

nonconsensual pornographic images should be evaluated, and should the proceeding be *ex parte*?

The *Dendrite* court appeared to prefer a probable cause standard. More precisely, it first describes probable cause as "a non-technical, flexible concept that does not require rigid, 'technical demands for specificity and precision'"[224] and goes on to note that "the District Court envisioned this four-part test to act as a flexible, non-technical, fact-sensitive mechanism for courts to use as a means of ensuring that plaintiffs do not use discovery procedures to ascertain the identities of unknown defendants in order to harass, intimidate or silence critics in the public forum opportunities presented by the Internet."[225] In a criminal case, that would be the standard for any search warrants, so we suggest explicitly adopting it here, even for civil cases.

Both the *Dendrite* and *Cahill* courts prefer that deanonymization requests not be *ex parte*. The *Dendrite* court quoted *Columbia Ins. Co. v. seescandy.com*[226] as saying "The requirement that the government show probable cause is, in part, a protection against the misuse of *ex parte* procedures to invade the privacy of one who has done no wrong."[227] The *Cahill* court spoke even more strongly: "When First Amendment interests are at stake we disfavor *ex parte* discovery requests that afford the plaintiff the important form of relief that comes from unmasking an anonymous defendant."[228] In fact, the entire point of notification to the defendant is to permit opposition at this stage of the proceedings.[229] This may, of course, be subject to delay in criminal cases.[230] In our scheme, however, the parties in the first two steps do not know the user's identity and hence cannot notify them. We thus suggest a probable cause standard and the opportunity for an adversarial process.

Not all courts endorse such a strict standard. The Ninth Circuit held that a grand jury investigating possible crimes did have a right to unmask users ' anonymity: "Any incidental infringement on Glassdoor's users' First Amendment rights is no more drastic than necessary to vindicate those

---

[224] *Dendrite*, 775 A.2d at 770. (Internal citation omitted)

[225] *Id.* at 771.

[226] 185 F.R.D. 573, 579–580 (N.D. Cal. 1999).

[227] *Dendrite*, 775 A.2d at 770.

[228] *Cahill*, 884 A.2d at 461.

[229] Determining that an image is nonconsensual pornography is not simple. Apart from issues of consent and actually matching the image to an actual complainant, there can be serious technical and legal issues involving what are known as "deep fakes"; *see generally* Kate Kobriger et al., *Out of Our Depth with Deep Fakes: How the Law Fails Victims of Deep Fake Nonconsensual Pornography*, 28 RICH. J. L. & TECH 204 (Aug. 2021).

[230] Fed. R. Crim. P. §41(f)(3).

compelling interests."[231] While this ruling is not obviously correct,[232] it underscores the need for statutory clarity. We are, after all, proposing a scheme that has the potential for broad infringement on anonymity.

The other issue we identified at the start of this section is how to ensure that the three parties necessary for deanonymization—the identity provider, the certificate authority, and the deanonymization agent—actually cooperate with legitimate requests. The legal issues are more straightforward, though implementation may be complex. We address them in the following part.

V. PROPOSED LEGAL CHANGES

A. Section 230

As Professor Citron has noted, a "healthy dose of humility is essential as we consider Section 230 reform."[233] The one major amendment to it, FOSTA, has created many new problems while likely not achieving its intended goals.[234] Therefore, and as noted in Part II.C, *supra*, we confine our change to the heart of Section 230 to one aspect of the changes proposed by Professor Citron. She suggests conditioning immunity on platforms taking five steps:[235]

> First, platforms should give individuals a way to report intimate privacy violations, cyber stalking, or cyber harassment. Second, they should have processes that enable them to address those reports. Third, they should endeavor to prevent intimate privacy violations, cyber stalking, or cyber harassment from recurring on their services. Fourth, platforms should be subject to certain minimum logging requirements so that individuals who sue users for online abuse can get access to the information needed to identify their abusers and prove their case in court. Fifth, platforms should remove, delete, or otherwise make unavailable intimate images, real or fake, that have been posted or shared without the subject's consent.

Our proposal addresses her fourth point: providing enough information to allow identification of abusers.

Our motive for endorsing this change is to provide an incentive for web sites to adopt our scheme: it would satisfy the logging requirement she

---

[231] *United States v. Glassdoor, Inc. (In re Grand Jury Subpoena)*, 875 F.3d 1179, 1191 (Court of Appeals 2017).

[232] Arbatman & Villasenor, *supra* note 67, at 114–116.

[233] *See* Citron, *supra* note 14, at 29.

[234] Jurecic, *supra* note 136. *See also* Citron, *supra* note 14, at §II.B.

[235] Citron, *supra* note 14, at 40–41.

proposed,[236] without making assumptions about number or type of devices employed by the offender, or how they are connected to the Internet.

The usual way starting point for identifying someone who has posted content begins with their IP address.[237] Professor Kosseff notes several problems with this:

> Plaintiffs face a few significant hurdles with this process. First, not all websites keep logs of IP addresses, and some sites that host particularly controversial user content have been known to avoid recording this data. Second, if a website, ISP, or anonymous user challenges a subpoena for identifying information, courts apply complex First Amendment balancing tests to determine whether they should enforce the subpoena. And third, even if a court enforces the subpoenas, the plaintiff will have only the name and contact information of the subscriber to the Internet connection from where the post originated. If, for instance, the user posted from a library or coffee shop, this information will be of little use in identifying the poster.[238]

There are additional problems. Pornographic pictures are often downloaded from one site and uploaded again to another. Tracing via IP address may catch subsequent uploaders, rather than the original one, but they are not the ones who should be liable; indeed, they may not know that the pictures were posted without consent. Privacy-preserving technologies such as Tor can make IP address-tracing useless.[239] Connections via cellular networks use a technology known as "network address translation," which lets multiple computers share a single IP address;[240] without going into

---

[236] *Id.* Internal citations omitted.

[237] An IP address is the Internet analog of a phone number; it is how computers on the net are actually identified by the underlying infrastructure. *See* J. Postel, *Internet Protocol*, No. 791 (Sep. 1981) or any networking textbook, *e.g.,* ANDREW S. TANENBAUM & DAVID J. WETHERALL, COMPUTER NETWORKS (5th ed. 2010). Common computer names, e.g., www.supremecourt.gov, are translated to IP addresses by a piece of Internet infrastructure known as the Domain Name System (DNS); *see* P. V. Mockapetris, *Domain Names—Concepts and Facilities*, No. 1034 (Nov. 1987).

[238] KOSSEFF, *supra* note 128, at 221.

[239] Dingledine et al., *supra* note 61.

[240] The Internet ran out of IP addresses long ago. Network address translation is a way for computers to share IP addresses when connecting to, e.g., web sites (*see, e.g.,* P. Srisuresh & K. Egevang, *Traditional IP Network Address Translator (Traditional NAT)*, No. RFC 3022 (Jan. 2001). Network address translation is employed by all home routers and most public or quasi-public Internet access sites, e.g., hotels, coffee shops, libraries, etc.

technical details, this sharing makes tracing via IP address even more difficult.[241]

By contrast, our scheme identifies a *person*, not a connection. Furthermore, per the requirements for identity providers, that identity is strongly established. While there may still be some ambiguity—nominally personalized credentials may have been uploaded to a computer shared by family members or domestic partners—this would narrow the pool of suspects to the point where traditional investigative techniques would almost certainly identify the offender, and would provide strong evidence for use in court.

We do, however, need to take this further. While Professor Citron suggests rulemaking by the Federal Trade Commission or some other agency would be needed,[242] and while such regulations could almost certainly accept our scheme, that alone is not sufficient. As discussed in Part IV, *supra*, we feel that statutory protections against abusive deanonymization should be enacted.

The first protection is, of course, technical: as described in Part IV, *supra*, three separate parties, the identity provider, the certificate authority, and the deanonymization agent, must cooperate in order to accomplish deanonymization. Some design alternatives could have reduced this to two parties; we rejected those, precisely to avoid the danger of too easy collusion or subornation.

We next suggest a probable cause standard of review. Our scheme inherently implicates the constitutional right to anonymous communication; this should not be done lightly, even if there is an allegation of criminal conduct. This is obviously even more true for civil suits.

It can, of course, be argued that the third party doctrine should apply, and that the necessary information should be obtainable by an ordinary subpoena. We disagree. Our scheme is intended to provide a proactive means of identification, even in the absence of any wrongdoing. We thus find it necessary to protect privacy to the greatest extent possible until a clear link to illegal behavior has been shown. Two of the most important Supreme Court decisions on anonymity link it to a right to privacy.[243] We

---

[241] Some additional details about network address translators are given in Steven M. Bellovin et al., *It's Too Complicated: How the Internet Upends Katz, Smith, and Electronic Surveillance Law*, 30 HARVARD JOURNAL OF LAW AND TECHNOLOGY 1, 49 (Autumn 2016).

[242] Citron, *supra* note 14, at 41.

[243] *NAACP v. Alabama ex rel. Patterson*, 357 U.S. 449, 462 (1958) ("This Court has recognized the vital relationship between freedom to associate and privacy in one's associations."); *McIntyre v. Ohio Elections Commission*, 514 U.S. 334, 341 (1995) ("The

thus follow Justice Harlan's observation in his concurrence in *Katz v. United States*: "that the invasion of a constitutionally protected area by federal authorities is, as the Court has long held, presumptively unreasonable in the absence of a search warrant."[244] Search warrants, of course, require a showing of probable cause.

We also suggest giving each of these three parties the statutory right to challenge their part in answering deanonymization requests.[245] That is, they should each have the opportunity to see the alleged non-consensual pornography and challenge the request in court if they felt it was abusive, e.g., if it did not show pornography at all or if there was otherwise no probable cause.[246] This must be done carefully. The obvious court to hear such cases would be the one that issued the search warrant in the first place, but that would mean that the same judge who issued the order based on their perception of probable cause would then rule on any challenges to it. This is regrettable but will often be unavoidable.

To the extent feasible, we also suggest enshrining parts of the *Dendrite* test into the statute.[247] However, we cannot do this literally. The first prong of the *Dendrite* test is notification of the purported abuser; only the identity provider can do that. Under certain circumstances, it may be possible for the certificate authority to do so, too, but these circumstances are sufficiently unusual that it probably unwise to codify them into statute.[248] The second, third, and fourth prongs are all subsumed by the probable cause requirement: if the images in question are intimate, and are plausibly alleged to be non-consensual by the victim,[249] the activity is clearly actionable, would survive a motion to dismiss for lack of a *prima facie*

---

decision in favor of anonymity may be motivated by fear of economic or official retaliation, by concern about social ostracism, or merely by a desire to preserve as much of one's privacy as possible.").

[244] *Katz v. United States*, 389 U.S. 347, 361 (1967) (Harlan, J., concurring).

[245] We discuss in Part B, *infra*, if these parties should be obligated to challenge such requests.

[246] *See* Part II.D, *supra*.

[247] *Dendrite Int'l, Inc. v. Doe No. 3*, 775 A.2d 756, 760–61 (N.J. App. Div. 2001).

[248] The circumstances have to do with renewal of an expiring certificate and how that would be handled by implementations. Certificates would typically have a long lifetime, measured in months or years, so the situation would not arise frequently. Furthermore, it is quite possible, and arguably better, if no certificates were ever renewed, as opposed to fresh ones being issued. In that case, there would be no possibility of notification by the CA.

[249] In a criminal case, the act of reporting the incident itself would generally carry criminal penalties if done falsely; *see, e.g.,* New York Penal Code § 240.50. In a civil case, an affidavit by the complainant would likely be necessary.

cause of action, and obviously cannot proceed without identification of the apparent perpetrator.

To summarize: we believe that our scheme is a better way to satisfy Professor Citron's fourth point than simple IP address logging would be. We also suggest that the first *Dendrite* prong should be imposed on the identity provider. While these could probably be established by regulation, we also feel that probable cause should be required for deanonymization and that the three parties in our system should have standing to challenge the order. These would most likely require statutory provisions.

B. Trusted Parties

The three parties necessary to implement our scheme are the identity provider, the certificate authority, and the deanonymization agent. They have certain crucial properties that must be enforced by regulation.

The simplest assumption is that any company subject to the jurisdiction of American courts will comply with valid legal processes. This leads to our first requirement: that all three parties be located within the United States, and hence subject to the authority of American courts. There is a minor issue, in that most cases will involve out of state parties, but the *Uniform Act to Secure the Attendance of Witnesses from Without a State in Criminal Proceedings*[250] can be used for criminal cases or the *Uniform Interstate Depositions and Discovery Act*[251] for civil cases.

Possibly, foreign companies might be eligible via a U.S. subsidiary. A foreign company's acquisition of a U.S. company filling one of these roles is the most likely cause of such situations; suitable conditions would have to be imposed by whatever regulatory bodies approved the acquisition, e.g., the Committee on Foreign Investment in the United States.[252] A purely foreign company is more problematic; though there are often mutual legal assistance treaties that could be used, the procedures are cumbersome and time-consuming.[253] It is unclear if the CLOUD Act would suffice.[254]

---

[250] National Conference of Commissioners on Uniform State Laws, *Uniform Act to Secure the Attendance of Witnesses from Without a State in Criminal Proceedings* (Aug. 1936).

[251] National Conference of Commissioners on Uniform State Laws, *Uniform Interstate Depositions and Discovery Act* (Aug. 2007).

[252] *See* https://home.treasury.gov/policy-issues/international/the-committee-on-foreign-investment-in-the-united-states-cfius.

[253] *See, e.g.,* Andrew K. Woods, *Data Beyond Borders: Mutual Legal Assistance in the Internet Era* (Global Network Initiative Jan. 2015) ("significant delays leave many law enforcement agents with the sense that the MLA process is a waste of time").

[254] Clarifying Lawful Overseas Use of Data Act, 115 P.L. 141, Division V (2018).

It is also important that the three parties be independent of each other. Our goal is to introduce enough friction in the process to discourage casual collusion: breaking anonymity should be hard. The implication here is that a single company should not be able to fulfill two or even all three roles. More precisely, though one company may have units that fulfill more than one role, more than one cannot be part of any single attestation. Although enforcement of this restriction may not be trivial, there are some technical features that should help. Identity Providers must know the public key of Deanonymization Agents, since the deanonymization string must be encrypted with it. Furthermore, the public keys of all legitimate CAs must be available to everyone, since arbitrary web sites need it to verify the certificates embedded in uploaded images.

Finally, since we wish to at least allow[255] for adversarial proceedings for deanonymization requests, geographic diversity is useful; if all three are in one small geographic area, the odds on a single magistrate or judge handling all three requests and thus granting the orders might be unpleasantly high.[256]

There is another, more subtle issue: all three parties must be honest.[257] The identity provider needs to verify user identities using criteria similar to those for notary publics; the certificate authority must maintain proper records, etc.

All of this leads us to the same conclusion: the qualifications for these three entities should be established by statute or regulation. Since we assume the need for a new statute in any event,[258] the necessary requirements can be enacted at the same time.

---

[255] As noted, this is a complex question discussed in Section V.C, *infra*.

[256] This issue has often been raised by Prof. Stephen I. Vladeck. *See, e.g., Amicus brief, Stephen I. Vladeck, United States et. al v. Texas and Louisiana*, 2022 WL 2466786, 5–6 (U.S.) ("In its recent lawsuits challenging federal policies, Texas has consistently exploited this situation, filing exclusively in those small divisions where it can all but guarantee which judge will hear its case. This case exemplifies this practice of Texas hand-selecting its judges.") and Steve Vladeck, *Texas Judge's Covid Mandate Ruling Exposes Federal "judge-Shopping" Problem*, MSNBC (Jan. 11, 2022), https://www.msnbc.com/opinion/texas-judge-s-covid-mandate-ruling-exposes-federal-judge-shopping-n1287324 ("The answer is "no"—it's savvy litigants of all partisan stripes, including the state of Texas, taking advantage of a little-known quirk in federal procedure in which, by filing in the right court, they can literally choose the judge who is going to hear their case.") *But see* the discussion in Part A, *supra*.

[257] In computer science terminology, they are assumed to be "honest but curious" *see. e.g.,* 2 GOLDREICH, *supra* note 191, at 603.

[258] *See* Part V, *infra*.

A closely related issue is who should select the parties and who should pay for them. It is important to keep performance incentives aligned with payments. Funding sources must be found for all three parties. In fact, there should be several of each, to provide competition and perhaps redundancy should one fail.

Competition is, of course, seen as a good thing in the American economy, but to cite it here raises the question of what these entities would compete on. Price is an obvious answer, though as we shall see it is not entirely satisfactory. Service is a second answer, though what service means is different for each of the three. Finally, there is the question of proxy representation in response to legal processes: will the certificate authority or the deanonymization agent vigorously protect users 'right to anonymity in court? Note that in this last case, the service interests of end users and those of law enforcement conflict. We heed the adage "Follow the money."[259]

This analysis suggests that deanonymization agents should not be paid by users: a miscreant user does not wish to be identified at all, let alone promptly, but if they were paying the agent the agent would be inclined to protect their interests. That said, users (including, of course, innocent ones) have a right to anonymity; some may wish for to pay deanonymization agents who will protect their interest in anonymity by contesting the request in court. All that said, it would be reasonable for parties requesting deanonymization, typically law enforcement agencies, to pay a per-request fee for the service;[260] while that might not pay for the necessary infrastructure, it could cover the marginal cost of each request.

The certificate authority is easiest to analyze. Except when handling deanonymization requests, there is no direct human interaction with it. As explained in Part III, *supra*, it is contacted automatically by a browser extension. The user does not directly engage the certificate authority. As such, it would make sense for it to be paid by the identity provider.

---

[259] The phrase is commonly associated with Carl Bernstein and Bob Woodward's classic book on Watergate, ALL THE PRESIDENT'S MEN (Simon and Schuster 1974); however, it did not appear there. Instead, it was in the movie version, Alan J. Pakula, *All the President's Men* (1976). The first use, though, was in a Senate confirmation hearing which, ironically, focused on Watergate issues. *Nomination of Earl J. Silbert, of the District of Columbia, to Be United States Attorney for the District of Columbia, 93rd Congress, S. Comm. on the Judiciary* 399 (Jun. 1974) (Statement of witness Henry Petersen, "I would say, 'Follow the money, Earl, because that's where it's going to be.'")

[260] Such a requirement is not unusual. The Wiretap Act provides for compensation to ISPs, landlords, etc., for their assistance in collecting data (18 U.S.C. § 2518(4)(e)). Similarly, the *Communications Assistance to Law Enforcement Act* authorized spending half a billion dollars to add the necessary functionality to the phone system (103 P.L. 414 §110, 1994).

The identity provider poses the most difficult questions. Users select them directly, and hence have the opportunity, and perhaps the obligation, to pay them directly. Furthermore, since for purely technical reasons the identity provider must know which certificate authority and which deanonymization agent are to be used for this person, a user could opt for an identity provider that offered a suitable choice. This could be good or bad—the user could select a certificate authority and a deanonymization agent that would act as their proxies to protect their anonymity, or they could choose one whose cooperation with legitimate requests was minimal.[261] In other words, "follow the money" cuts both ways.

It is clear, though, that legal representation would be a non-trivial expense. One online notary charges $25 per operation just for identity verification;[262] identity providers are more complex, since users 'browsers must interact with them continuously, though all interactions save for the initial registration are automated and rapid. The potential for legal representation would be an additional expense, with payment passed on from the identity provider to the other two entities. The expense is not a novel issue; as a classic *New Yorker* cartoon says, "You have a pretty good case, Mr. Pitkin. How much justice can you afford?"[263] Nevertheless, it exacerbates the social issues discussed below. Furthermore, it is unclear how well these systems actually work. One, ID.me, has been criticized by a number of members of Congress.[264]

Another alternative is to have web sites that rely on our scheme for Section 230 protection pay a "tax" that would support the various pieces of the infrastructure. This makes technical sense, since a participating web site must know which certificate authorities it should trust.[265] Those who wished might pay a surcharge to cover legal representation by the downstream

---

[261] It is, in fact, technically feasible for an identity provider to choose randomly among several certificate authorities and deanonymization agents.

[262] *See* https://www.notarize.com/pricing.

[263] J.B. Handelsman, *Cartoon*, NEW YORKER, Dec. 24, 1973, at 52.

[264] Ron Wyden et al., *Deceptive Statements Made by ID.Me* (May 2022) ("[T]he use of one-to-many recognition means that millions of innocent people will have their photographs endlessly queried as part of a digital "line up." Not only does this violate individuals' privacy, but the inevitable false matches associated with one-to-many recognition can result in applicants being wrongly denied desperately-needed services for weeks or even months as they try to get their case reviewed. This risk is especially acute for people of color: NIST's Facial Recognition Vendor Test found that many facial recognition algorithms have rates of false matches that are as much as 100 times higher for individuals from countries in West Africa, East Africa and East Asia than for individuals from Eastern European countries.")

[265] The reasoning here is technical and there are ways around it, albeit at the cost of more centralization.

entities. On the other hand, a taxation scheme poses its own difficulties: an identity provider that offers more personalized service, or one that offers in-person service in a large, rural area would inherently cost more; allocating the payments would be a non-trivial problem.

To summarize: since the identity provider must select the certificate authority and the deanonymization agent, it would make sense for the provider to pay the other two, though the latter could collect additional fees. However, the cost of the service may be prohibitively high for some people, in addition to the documentation issues discussed earlier. Web sites benefit the most, so they would be one source of funding, if a suitable allocation formula can be devised. This issue does require further study.

Taken together, all of these points suggest the need for licensing of the various parties. License-implementing regulations would assure that the parties had the technical capabilities to respond to lawful requests and had sufficient legal resources available to challenge apparently invalid requests. Licensing—more precisely, license renewals—would provide a safeguard against improper mergers or takeovers, especially by offshore companies. Finally, it may be desirable to impose price controls, though that question demands a more in-depth economic analysis.

C.Economic and Social Issues

Although our scheme can protect privacy while providing accountability, there are other costs that must be borne. Chief among them is the requirement that would-be image posters have an arrangement with an identity provider that they would pay. This raises questions that do not have easy answers. Here, we outline the issues but do not provide definitive solutions.

The requirement for strong identity verification is the most troubling, since it brings in issues of racial and economic equity. As noted earlier, an identity provider requires the same level of assurance as a notary public. Notaries, though, typically want their clients to provide identifying documents; not everyone has them. This has been an issue in, *e.g.*, Voter ID laws, the burden of which can fall disproportionately on the elderly, the poor, those living in rural areas, etc.[266] Undocumented individuals, if they

[266] *Crawford v. Marion County Election Bd.*, 553 U.S. 181, 199 (2008) ("Both evidence in the record and facts of which we may take judicial notice, however, indicate that a somewhat heavier burden may be placed on a limited number of persons. They include elderly persons born out of State, who may have difficulty obtaining a birth certificate; persons who because of economic or other personal limitations may find it difficult either to

did not have local identification documents, could use identification documents from their original countries.[267] There are online identification and notary services, but these typically also require some sort of government-provided photo ID.[268] One common identification document, a bank statement, is only available to those with bank accounts, but poor people and racial minorities are less likely to have such accounts.[269] The U.S. Department of State will accept identification statements from witnesses for passport applications,[270] but even this can be problematic. Not only are disadvantaged people more likely to disproportionately know other disadvantaged people, the very desire to upload images may pose social questions: "What sorts of pictures do you want to upload, and why?" This latter question may be asked by family members, but it may also be asked by government agencies that are afraid of whistle-blowers.

There may even be legal bars to such a scheme if there are no feasible work-arounds available. In a New Mexico case about a law requiring use of,

---

secure a copy of their birth certificate or to assemble the other required documentation to obtain a state-issued identification; homeless persons; and persons with a religious objection to being photographed.") (Internal footnote omitted.) Justice Souter, in his dissent, takes it further: "The first set of burdens shown in these cases is the travel costs and fees necessary to get one of the limited variety of federal or state photo identifications needed to cast a regular ballot under the Voter ID Law. The travel is required for the personal visit to a license branch of the Indiana Bureau of Motor Vehicles (BMV), which is demanded of anyone applying for a driver's license or nondriver photo identification. The need to travel to a BMV branch will affect voters according to their circumstances, with the average person probably viewing it as nothing more than an inconvenience. Poor, old, and disabled voters who do not drive a car, however, may find the trip prohibitive, witness the fact that the BMV has far fewer license branches in each county than there are voting precincts.") (Internal footnotes and citations omitted.) *Id.* at 211–213 (Souter, J., dissenting).

[267] A number of states explicitly permit undocumented individuals to receive driver's licenses. New York passed such a law a few years ago (2019 N.Y. Laws, Ch. 37). *See also Kearns v. Cuomo*, 981 F.3d 200, 204 (2nd Cir. 2020) ("The REAL ID Act does not bar states from continuing to issue driver's licenses that do not comply with the Act… Notably, the REAL ID Act does not require states to verify the lawful status of applicants for noncompliant licenses.") (Internal citations omitted.)

[268] *See, e.g.,* Notarize.com ("Notarize uses identification verification technology to verify government-issued photo IDs and passports. Take a picture, answer a few questions, and we'll confirm your identity in seconds."), *available at* https://www.notarize.com/how-it-works, and ID.me ("ID.me uses facial recognition to match the user's selfie to their uploaded government ID."), *available at* https://www.id.me/business/identity-gateway.

[269] Paola Boel & Peter Zimmerman, *Unbanked in America: A Review of the Literature* (Federal Reserve Bank of Cleveland May 2022).

[270] *See* Form DS-71, *available at* https://www.reginfo.gov/public/do/DownloadDocument?objectID=41250301. The form states that the estimated burden is five minutes, but that does not include the time needed for the witness to appear before an authorized passport acceptance agent.

e.g., a credit card to prove that one is an adult and hence eligible to access adult content, a court wrote that "[r]equiring a credit card, debit account, adult access code, or adult personal identification number before providing access to speech on the Internet would bar many adults who lack such identification from access to information appropriate for them."[271] We are not proposing credit cards, but many of the barriers are the same. Furthermore, while that case was about access to information, our scheme is about creating content, which is far more clearly expressive and hence far more clearly protected by the First Amendment.

A possible workaround is what is known as "social authentication."[272] In social authentication, previously enrolled users vouch for someone else, either as a new user or to deal with lost passwords or other credentials. This notion is in fact compatible with other government authentication requirements. For example, passport applicants may prove their identity via affidavits from other people: "The applicant must establish his or her identity by… other identifying evidence which may include an affidavit of an identifying witness,"[273] though there might need to be other requirements established, as indeed there are for passports.[274] Similarly, Massachusetts will permit approved social services agencies to file affidavits of state residence on behalf of homeless individuals whom they are aiding.[275]

It is important to understand what such affidavits do and do not do. They do not make the affiant responsible for content uploaded by others. They do not make the affiant responsible for real-time, continuing knowledge of the whereabouts of the person whose identity they are vouching for. Rather, they are taking the place of traditional identity documents: "I attest that this person is known to me as XXX and resides at YYY."

To sum up: the poor, minorities, and some others will have difficulty obtaining the credentials necessary to use our scheme. This could block

---

[271] *ACLU v. Johnson*, 4 F. Supp. 2d 1029, 1032 (D. N.M. 1998), *aff'd ACLU v. Johnson*, 194 F.3d 1149 (10th Cir.).

[272] Hyoungshick Kim et al., *Social Authentication: Harder than It Looks*, Proceedings of Financial Cryptography and Data Security (2012).

[273] Identity of applicant, 22 C.F.R. § 51.23(b) (2023).

[274] *Id.* § 51.23(c) ("The Department may require such additional evidence of identity as it deems necessary.")

[275] https://www.mass.gov/guides/rmv-real-id-info-center ("We also understand that certain populations may struggle with obtaining the required documents to get or renew a credential. The RMV has created an affidavit that may be used for individuals who are unable to prove Massachusetts residency while actively receiving services from an official institution… [T]his affidavit may be accepted as ***one proof*** of Massachusetts residency for a **Massachusetts identification card only**.") (Emphasis in the original.)

many of them from uploading images to popular sites that require it. That said, given that this is a larger social problem, there are larger efforts to address it.[276] If they succeed, or if other forms of identification are accepted, this will become a non-issue.

A somewhat-related issue concerns the web browser used to upload images: it must have the user's credentials in order to sign the images.[277] This may not always be the case, especially if the images are uploaded from a public computer, e.g., one in a library.[278] Such a device would not and should not have any per-user private data.[279] The solution would be to have the cameras themselves sign the photos in the appropriate fashion. This is especially easy for the most common kind of camera, the cell phone, since they contain web browsers and have Internet connectivity. It is somewhat more problematic for standalone cameras; while they can easily sign images,[280] they would need to do so with appropriate privacy-preserving credentials and that would require some sort of connectivity, either directly or via a link to a smartphone. A solution there awaits further technical work, though this should not be arduous: if the camera does not have Internet

[276] *See, e.g.,* Teresa Wiltz, *Without ID, Homeless Trapped in Vicious Cycle* (Pew Charitable Trusts May 2017)*, available at* https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2017/05/15/without-id-homeless-trapped-in-vicious-cycle.

[277] Most web browsers have a privacy-preserving feature called "incognito mode" windows, which do not retain or share identifying data such as cookies. To preserve that property, an incognito window in our scheme would use a separate store of certificates and would delete them frequently.

[278] People may opt to use such computers if they don't have computers of their own at home, which is especially common in poorer households; *see, e.g.,*Kendall Swenson & Robin Ghertner, *People in Low-Income Households Have Less Access to Internet Services* (Office of the Assistant Secretary for Planning & Evaluation, U.S. Department of Health & Human Services Apr. 2020). Alternatively, they may use public computers precisely to evade tracking.

[279] There may also be a problem if a user owns multiple devices, e.g., a phone, a laptop, a tablet, etc. In such a situation, the credentials must be copied to each such device. While this is not a trivial problem, technical solutions and workarounds for credential movement have been developed; *see, e.g.,* John S. Koh et al., *Encrypted Cloud Photo Storage Using Google Photo*, MobiSys 2021 (2021).

[280] The notion of cameras digitally signing images is not new; *see, e.g.,* J. Kelsey et al., *An Authenticated Camera*, Proceedings 12th Annual Computer Security Applications Conference 24 (1996). At least two manufacturers, Canon and Nikon, actually marketed such devices, though due to incorrect implementations their schemes have been cracked. *See* Olga Koksharova, *Canon Cannot or Mustn't Provide Image Validation Feature?*, ELCOMSOFT BLOG (Nov. 30, 2010), https://blog.elcomsoft.com/2010/11/canon-cannot-or-mustnt-provide-image-validation-feature/ and Vladimir Katalov, *Nikon Image Authentication System: Compromised*, ELCOMSOFT BLOG (Apr. 28, 2011), https://blog.elcomsoft.com/2011/04/nikon-image-authentication-system-compromised/.

connectivity of its own, its images would be uploaded via a web browser; if it does, our scheme could be implemented directly.[281]

D.Security Analysis

Although this is not a security paper, it is reasonable to ask what the security risks are. More precisely, what are the risks if some component is hacked or corrupt, and what can be done about that?

Any security analysis has to start by asking "What are you trying to protect, and against whom?"[282] In this situation, the first part is easier to answer: we are trying to ensure that the entire process gives the right answers. That is, it should neither fail to identify someone who did upload NCP, nor should falsely identify someone who did not upload a specific image. The system should also protect the privacy and anonymity of users except under the conditions outlined here. The threats are technical, e.g., a computer has been hacked or the code is incorrect, or procedural: parties collude when they shouldn't, or decline to cooperate with valid court orders.

The latter is easier to analyze. We designed our system to require cooperation of three independent parties precisely to minimize the risks from collusion or corruption. It is clear that at least on occasion, law enforcement personnel abuse their access to databases.[283] For that reason, we provide for judicial review of all requests, at every stage: a single

---

[281] Many modern cameras do have WiFi capability. *See, e.g.,* Using Your Nikon Camera's Built-in Wi-Fi (Jun. 2022), https://www.nikonusa.com/en/learn-and-explore/a/tips-and-techniques/using-your-nikon-cameras-built-in-wi-fi.html ("All Z series Mirrorless cameras, select Nikon DSLRs and many COOLPIX compact digital cameras have wireless connectivity built-in.").

[282] STEVEN M. BELLOVIN, THINKING SECURITY: STOPPING NEXT YEAR'S HACKERS 31 (Addison-Wesley 2016).

[283] *See, e.g.,* Dhruv Mehrotra, *ICE Records Reveal How Agents Abuse Access to Secret Data*, WIRED (Apr. 17, 2023), https://www.wired.com/story/ice-agent-database-abuse-records/ ("According to an agency disciplinary database that WIRED obtained through a public records request, ICE investigators found that the organization's agents likely queried sensitive databases on behalf of their friends and neighbors. They have been investigated for looking up information about ex-lovers and coworkers and have shared their login credentials with family members. In some cases, ICE found its agents leveraging confidential information to commit fraud or pass privileged information to criminals for money.") or *United States v. Valle*, 807 F.3d 508, 512–513 (2nd Cir. 2015) ("As an NYPD officer, Valle had access to the Omnixx Force Mobile ('OFM'), a computer program that allows officers to search various restricted databases, including the federal National Crime Information Center database, which contain sensitive information about individuals such as home addresses and dates of birth.") "Valle concedes that he violated the terms of his employment by putting his authorized computer access to personal use", *Id.* at 523.

corrupt officer or police force[284] cannot unilaterally deanonymize the contributor of an image.

Lack of cooperation with valid court orders would be dealt with through ordinary legal processes, including findings of contempt and possible seizure of the requisite databases. Per Section V.B, *supra*, the parties whose cooperation is necessary for deanonymization are all regulated; at a minimum, non-cooperation could result in a loss of their license to do business.

Technical flaws are harder to analyze. We can be confident that a single point of failure, i.e., one of the parties being hacked, will not result in deanonymization; again, all three must fail. There is more risk of a failure to identify someone correctly due to ordinary software bugs, which are of course legion. There are no perfect solutions known, though there are heuristics that can guide software selection.[285]

A more serious problem is misidentification: returning the identity of the wrong person, since that can lead to an innocent person being convicted of a crime. Again, there are no perfect solutions possible. One author of this work has suggested elsewhere that defendants have a constitutional right of access to "evidentiary" source code.[286] This will permit adversarial audits to try to find bugs. This principle has already been accepted by at least one state appellate court.[287]

The strongest part of our solution is the cryptographic mechanisms involved. The most complex piece, the base Camenisch-Lysyanskaya protocol,[288] contains formal mathematical proofs of correctness, passed peer review, and has stood the test of time. Our additional mechanisms are sufficiently straightforward that they are very likely to be correct, too.[289]

We must also consider who the likely attackers are. In an earlier work, one of the authors of this article classified attackers along two axes, skill level and how specifically a particular entity was targeted.[290] Here, the

---

[284] *See supra* note 86.

[285] *See, e.g.,* BELLOVIN, *supra* note 276, ch. 12. While that deals with security flaws in software, the same criteria can be used to assess the probability of ordinary flaws.

[286] Steven M. Bellovin et al., *Seeking the Source: Criminal Defendants' Constitutional Right to Source Code*, 17 OHIO STATE TECHNOLOGY LAW JOURNAL (Dec. 2020), https://moritzlaw.osu.edu/ostlj/2020/12/22/seeking-the-source-criminal-defendants-constitutional-right-to-source-code/.

[287] *State v. Pickett*, 466 N.J. Super. 270 (N.J. Super. Ct. App. Div. 2021).

[288] Camenisch & Lysyanskaya, *supra* note 28.

[289] We in fact plan to submit the companion technical paper for publication, at which point it, too, will undergo peer review.

[290] BELLOVIN, *supra* note 276, at 34.

specific targets are known, the three parties in the total system; the question is motivation. That in turn divides into inculpatory and exculpatory hacking.

Clearly, the parties with the most incentive to do exculpatory hacking are people who are actually guilty. While there is no reason to believe that such people possess unusual computer hacking skills, they could presumably hire someone who did possess them.

The easiest attack is to hack someone else's computer, steal their credentials, and use those to post NCP imagery. That, of course, would have to be done in advance, and if the security features of modern computers are used, stealing credentials would be extremely difficult.[291] Accordingly, a more likely target for a hacker would one of the three parties in our system. For complex technical reasons, the best parties to hack for exculpatory reasons are the identity provider and the certificate authority.[292] The most feasible attack would involve changing their databases to point to someone else, thereby inculpating them.

A way to inculpate someone else would involve a procedural hack: presenting false credentials to the identity provider. These, however, are held to approximately the same standards as notaries, and is thus a risk that the legal system takes every day.

The remaining question is who would launch these attacks, other than a party wishing to escape liability. One answer is a deliberate desire to frame someone, for revenge or perhaps to influence an election. The latter is a serious concern, since it might be carried out by an extremely capable intelligence agency. But there have been instances of planted computer evidence done for no apparent reason.[293]

---

[291] Windows PCs contain a Trusted Platform Module chip; *see, e.g.,* Tom Brant, *What is a TPM, and Why Do I Need One for Windows 11?*, PC MAGAZINE (Sep. 24, 2022), https://www.pcmag.com/news/what-is-a-tpm-and-why-do-i-need-one-for-windows-11. Newer Apple computers use a T2 chip, which provides essentially the same functions, *Id.*; *see also Apple T2 Security Chip: Security Overview*, https://www.apple.com/jp/mac/docs/Apple_T2_Security_Chip_Overview.pdf (last visited Apr. 21, 2023). Both chips store cryptographic keys and carry out cryptographic operations; however, the cryptographic keys are not readable by the host computer, even if it is hacked. This provides very strong protection against credential theft.

[292] Hacks of the other two parties are immediately detectable, and the proper data is likely recoverable from backup media. For details, see the security analysis in the companion technical paper, Gorman et al., *supra* note 175.

[293] *See, e.g.,* getreading, *Program Put Child Porn Pictures on My PC*, BERKSHIRE LIVE (Apr. 16, 2003), https://www.getreading.co.uk/news/local-news/program-put-child-porn-pics-4271386.

E.Mission Creep

Another issue is what is often called mission creep: using a mechanism created for one purpose for another. We have proposed a system intended for use in cases of non-consensual pornography; might it be used for other crimes as well, e.g., child sexual abuse material or terrorism? The example of the Wiretap Act is cautionary. The original version of the law restricted wiretaps to seven classes of crimes.[294] By contrast, today's law lists 21 classes, adding crimes involving things like sale of immigration papers, damaging gas or oil pipelines, illegal restraint of trade, and more.[295]

There does not appear to be an easy legal or technical solution to this problem. Hypothetically, one could posit a statutory provision barring use of evidence collected by credentials issued before such an amendment to this law was passed, but of course Congress could repeal that provision, too. Possibly, a policy that evidence collected via these mechanisms could only be used for crimes already specified when the infringing actions took place could be incorporated into the Federal Rules of Criminal Procedure, since changes to it must originate with the Supreme Court, not Congress.[296] This does not prevent mission creep, but it does slow it down; more importantly, it requires the cooperation of a separate branch of government. But this protection poses the same obstacle as its benefit: the judiciary would have to adopt it independent of Congressional action.

There might be a technical solution. Assume that most instances of non-consensual pornography are discovered within a fairly short time—and this an assumption; we know of no data on the subject—the cryptographic certificates used to sign images, and their associated database entries, could be deleted after that interval.[297] The expiration would be transparent to the user save for a short delay when posting an image; the browser extension would simply obtain a new certificate from the CA. The database entry contains the deanonymization string; without it, there would be no way to learn who uploaded the image, and hence no ability to abuse the deanonymization facility. Of course, this also assumes that abuses would primarily occur after that interval.

---

[294] Omnibus Crime Control and Safe Streets Act of 1968, Pub. L. No. 90-351, Title III, 82 Stat. 197 §2516.

[295] 18 U.S.C. § 2516(1)(a-u).

[296] 28 U.S.C. § 2072 (a) ("The Supreme Court shall have the power to prescribe general rules of practice and procedure and rules of evidence for cases in the United States district courts (including proceedings before magistrates [magistrate judges] thereof) and courts of appeals.")

[297] As noted, certificates generally contain an expiration date; *see* Part III.B, *supra*.

Most likely, though, the answer is procedural: any new additional uses must be justified by the same sort of exacting scrutiny analysis we have provided here. How well that would work would, of course, depends on the philosophical leanings of Congress and the courts.

VI. CONCLUSIONS

There is little doubt that this scheme can work from a technical perspective. Whether it should be deployed is a harder question.

The first issue is economic. There is no doubt that there is noticeable expense involved in this proposal: someone would have to fund the identity providers, the certificate authorities, and the deanonymization agents. The ordinary operation of the CAs is likely inexpensive and would be completely automated; the expense of responding to legal requests might be considerable. Still, given the large number of people who upload benign images, amortizing the cost of legal responses is likely feasible, though a more detailed economic analysis should be performed. Of course, a base funding source must be identified. For the deanonymization agent, the legal expenses would be by far the largest cost; as noted, though, this could be billed back to law enforcement.

The cost of the identity provider is the most problematic, since it is one the most likely to be imposed directly on users. Here, the costs are running a highly available data service—which is doable, and whose cost can be amortized—the legal expenses, and (most critically) the cost of identity verification. That would fall squarely on the users, as discussed in Part V.C, *supra*.

Cross-border issues present another challenge: what if images are initially uploaded to foreign sites, ones that do not participate in this scheme? There is no simple answer, save to note that the situation would be no worse than it is today with unsigned images. We do note that most of the large tech companies are American, so this may not be a huge issue. It is also quite plausible that other countries would adopt similar schemes. An analysis of how cross-border deanonymization should work would be interesting but well beyond the scope of this article.

The biggest problem, though, is the possibility of mission creep. The history of the Wiretap Act is not encouraging. The notion that wiretap warrants should be limited to serious crimes goes back to at least 1952.[298] In fact, even when the original law was adopted in 1968 the list of offenses

[298] Alan F. Westin, *The Wire-Tapping Problem: An Analysis and a Legislative Proposal*, 52 COLUMBIA LAW REVIEW 165, 202–203 (Columbia Law Review Association, Inc. 1952). Westin noted just how despised wiretapping was; *id.* at 166.

was seen by some as too great. During the floor debate on the bill, Rep. Robert Eckhardt said, "The provisions of the bill contained in Title III are lamentable encroachments on privacy and on constitutional rights. They permit wiretapping to seek evidence concerning a great variety of offenses which may be charged against a citizen—and we must remember that the charges may not ultimately be borne out, but in seeking to do so the citizen's privacy is shamefully invaded."[299] A Congressionally mandated review a few years later showed that this was still a concern: "A substantial minority of the Commission… found that… court-authorized surveillance… even under the authorization and supervision of a court has resulted in substantial invasions of individual privacy"[300] and that "that the list of Federal offenses for which a wiretap order can be sought should be reduced rather than expanded."[301] Despite all that, and as noted, the list of crimes has grown considerably.

The mission creep issue deserves deep consideration. Non-consensual pornography is a very serious problem, but so is the loss of anonymity online.

[299] Cong. Recd., 90th Cong., June 6, 1968, p. 16274, *available at* https://en.wikipedia.org/wiki/90th_United_States_Congress.

[300] *Report of the National Commission for the Review of Federal and State Laws Relating to Wiretapping and Electronic Surveillance* xiii (Apr. 1976), *available at* https://www.ojp.gov/pdffiles1/Digitization/39007NCJRS.pdf.

[301] *Id.* at 5.