Research Statement

Sarah Scheffler

My research at the forefront of **applied cryptography** and **policy** identifies the societal shortcomings of existing socio-technical systems that involve encryption and builds cryptographic protocols to address those shortcomings. My research record shows not only my aptitude with the standard applied cryptography toolkit of building cryptographic protocols and crafting security proofs, but also my ability to synthesize a complex technical field with challenging policy tradeoffs, as well as experience with empirical legal scholarship. My work makes impactful contributions both intellectually and through direct engagement with policymakers.

My contributions fall into two main areas: First, my work on **encryption policy** combines strong policy and legal analysis with novel developments in cryptography. I show my strength in this joint analysis in two key areas: analyzing content moderation under encryption, and understanding reach and limits of court-mandated decryption orders. Second, I also advance the field of applied cryptography in a "purely" technical way, advancing the frontier of **zero-knowledge proofs** and bringing the cryptographic notion of secure multi-party computation to mapping problems in robotics.

1 Encryption policy

Over the last several decades, cryptography has grown from a specialized niche tool into a ubiquitous technology in all layers of our digital infrastructure. At the highest layer, *end-to-end encryption* (E2EE) brings confidentiality and authentication directly under the control of users—and since only the users have the keys, even the service providers cannot read or tamper with the plaintext content. E2EE reached mass adoption in the last decade across many domains, especially for chat (WhatsApp), video calling (optional E2EE in Zoom), full disk encryption (deployed by default on most mobile devices), and even web browsing (HTTPS hides the contents of web page requests from ISPs and other network infrastructure). This is a win for privacy and security for average people—but brings with it governance challenges, as encryption makes it harder to catch misbehaving users of these E2EE services.

My research over the last few years has provided key contributions in both cryptography and policy toward solving these governance challenges. My research on *content moderation under end-to-end encryption* directly seeks to understand and improve the possibilities for content moderation under E2EE, while making the systems as transparent and auditable as possible, so that users can be confident the content moderation tool has not been repurposed for mass surveillance. I will also discuss my work on *compelled decryption*, which seeks to understand when U.S. courts can and cannot order people to decrypt devices, a question which courts and legal scholars are struggling to answer with analogies.

Content Moderation under End-to-End Encryption

End-to-end encryption's core goal of providing confidentiality against the underlying service provider also disrupts many naive content moderation systems that provider might have performed to detect spam, malware, child safety violations, hate speech, or other problematic content. The year 2021 was a flashpoint in the global policy debate over content moderation in end-to-end encryption with the proposal of efficient hash matching systems [12, 45]. These systems would automatically alert the provider if plaintext content shared a hash with known Child Sexual Abuse Material (CSAM), while preserving the privacy of non-hash-matches. Most publicly, Apple's proposal [12] was criticized for creating risks to free speech, privacy, and security [2, 40, 56, 50, 30]. Apple delayed the most controversial system indefinitely [15], but governments took note of the possibility of proactively detecting CSAM, terrorist, or other content under encryption (e.g. [20, 52, 38, 21]). All stakeholders struggle to weigh the risks these systems pose to freedom of expression and

security, against tangible harms against children and disinformation targets and aim to settle on the frontier of the tradeoff. My research aims to tackle this problem from all sides, with a special focus on transparency mechanisms that roadblock the slope to censorship.

Establishing the foundation. In [SM23], I completed a Systematization of Knowledge (SoK) on content moderation for end-to-end encryption which was accepted at the Privacy Enhancing Technologies Symposium 2023, a top venue for the specific subject matter of privacy. I also presented an early version of this work at the DIMACS Computer Science and Law workshop on Content Moderation in May, 2022.

To do the first part of this work, I performed a massive literature analysis touching thousands of papers total and doing a deep dive into about 120 industry and academic proposals for content moderation under E2EE. My analysis unifies the privacy-preserving content moderation literature and includes not only the oft-discussed policy challenges of child safety and disinformation, but also topics that were nearly absent from the current debate, like corporate and parental monitoring of encrypted internet traffic.

In the second part, I provide much-needed contextualization of this existing technical research on content moderation in E2EE. I provide a general framework to analyze design choices in privacy-preserving content moderation that is useful for policymakers, service providers, technologists, and civil rights advocates alike. I go into detail on the current set of options for content moderation under encryption, including the level of privacy, detection mechanism, cryptographic tools used, security guarantees, efficiency, and any transparency properties the system offers. This work crystallizes the importance of a deep technical and policy understanding of content moderation in E2EE – there are already deployments of E2EE content moderation, and they do in fact suffer from misuse and lack of transparency (e.g. [57, 54, 53]).

Incorporating transparency and auditability. To that end, my research also builds technical transparency mechanisms for content moderation under E2EE in a paper [SKM23] conditionally accepted at IEEE Security & Privacy 2023, a tier-1 computer security venue. We begin with a detailed policy analysis of the issues at stake for U.S.-based deployers of end-to-end encrypted content moderation, and identify specific areas in which cryptographic protocols can improve the transparency and reliability of a content moderation system. We then provide constructions and implementations for each of our identified use cases for the specific case of Apple's controversial PSI system from 2021 [12].

To *build trust in the implementation* we create a system that enforces notification to users if their content was revealed to Apple, after a delay allowing the moderator to process the detection in some way (e.g. passing it to law enforcement). We implemented this system using the state-of-the-art malicious-secure authenticated garbling approach to multi-party computation of Wang et al. [62, 61]. To *build trust in the hashset* we provide two contributions. The first is a threshold signature scheme [4, 19] allowing child safety organizations to certify their part of the hash list in a publicly verifiable manner, while ensuring the list remains private to clients and robust against malicious attempts to alter the list. Second, we build a scheme that allows the central moderator to prove that specific hashes are *not* contained in the hashlist, providing credibility for claims that they are only using the moderation scheme for its intended purpose. To do so, we propose using a zero-knowledge proof of non-membership in a Cuckoo table of blinded hash values. Proving non-membership in a set could be accomplished by a negative accumulator [7, 8], but this generic approach is inefficient, and does not make use of the existing public information in our specific content moderation setting. Our approach shrinks the computation time and communication required by making use of existing public information. By combining homomorphic commitments [29, 55] with the classic proof of knowledge of discrete logarithm by Chaum et al. [18], we create a proof of non-membership whose size is dependent only on the security parameter. These interventions combined raise both the technical and normative bars to misusing a content moderation system for more censorious purposes.

Engaging with policymakers. As a part of this research agenda, I met with a group of engineers and decision-makers at Apple, including privacy chief Erik Neuenschwander, to provide feedback on the design of Apple's proposed content moderation systems. Separately, I have actively participated in formal discussions of the U.K.'s Safety Tech Challenge Fund [1], an initiative by the U.K. Home Office to build end-to-end encrypted systems that moderate content for child safety goals. I submitted formal written feedback [SMK22] during the Challenge Fund's evaluation process which argued for improved transparency, abuse resistance, and accuracy.

Future work. I have already begun two additional lines of work on this topic: First, I turn to one of the most pressing motivations for moderating content under encryption: to preserve the safety of children against various forms of sexual abuse. Although this topic is at the forefront of the debate over legislating access to plaintext in democratic governments (e.g. [52, 21, 36]), surprisingly little research has been conducted on the scope of the problem, and the extent to which encryption does or does not stymie child safety efforts in the U.S. In ongoing empirical legal research, I am investigating public dockets to examine the role encryption plays in U.S. Federal District Court prosecutions of 18 U.S.C. 2252 and 2260, which prohibit possession and production of CSAM and some related harms against children. Second, I am researching the possibilities for cryptographic improvement of the transparency and integrity of the terrorism-related hash-sharing database of the Global Internet Forum to Counter Terrorism, in collaboration with a member of their Technical Working Group.

Recent work shows that content moderation under end-to-end encryption is technically possible, and there is obvious governmental and business interest in deploying analysis of encrypted information. However, the threats to privacy, security, and free speech remain. With this in mind, the time to research transparency and auditability in these systems is *now*. My existing research on this topic, as well as my research on multi-party computation and zero-knowledge proofs, make me well-placed to conduct research at the forefront of this critical area.

Compelled Decryption

Several years of litigation have failed to answer an increasingly common question of U.S. law: *Can the government compel a device's owner to enter their password to decrypt their device?* The Fifth Amendment of the Constitution allows defendants to remain silent and refuse to perform specific types of compelled orders. However, it is not clear how this right against self-incrimination applies to decryption orders. The U.S. court system has struggled with not only competing state Supreme Court rulings [22, 23, 24, 59, 60] but also competing theories of the underlying legal doctrine [43, 58, 63] and the issue is expected to reach the federal Supreme Court soon. My research in this area has two goals: first, I seek to put the legal theories of compelled decryption on sound self-consistent footing. With that formalization in place, I then consider the ramifications of compelled decryption on cryptographic research and design.

Using cryptography to improve legal understanding. My research combines legal expertise with a technical understanding of cryptography, allowing direct evaluation of a question that legal scholars tend to rely on analogies for. My first paper on the topic [SV21][1] was published in Usenix Security 2021, another top-tier conference in computer security; it is one of very few interdisciplinary papers published at this venue and it formed the backbone of my Ph.D. dissertation. In this work, I created a formal model to determine whether an action is compellable under specific evidence based on a key legal principle that this form of compelled action should "add[] little or nothing to the sum total of the Government's information" [27]. This work has received praise from both legal and computer science audiences, including at the Privacy Law Scholars Conference 2021, Real World Crypto 2021, and the 2020 DIMACS Workshop on the Co-Development of Computer Science and Law.

Building upon that work, I also published a more recent work on this topic [CSV22][1] at the second ACM Symposium on Computer Science and Law (ACM CS/Law) in November, 2022. This newer work extends the computational framework, allowing us to analyze more complex scenarios like deniable encryption, using (potentially-randomized) methods in oracles and allowing for partial specification. This work revolved around a different key legal principle, that the government should not "rely on the truthtelling" of the respondent's implicit testimony [27]. Across both works, I have been invited to speak on this topic at twelve universities and institutions, including Stanford, UC Berkeley, MIT, Carnegie Mellon, and Microsoft Research.

Using legal understanding to guide cryptographic design. With a formal model in hand, in the second part of [SV21][1] I analyzed the consequences on real-world cryptosystems. Among other findings, the model showed that the secret inputs to multi-party computation were particularly vulnerable to being compelled in a way that did not apply to other cryptosystems. I thus considered the flip-side of the question:

[1] The author order of these papers follows an alphabetical convention and order does not represent contribution.

could I define and build a system without this weakness? I defined *FC-resilience* to capture this security property, and constructed an FC-resilient form of multi-party computation that I made available as open-source code [Scheffler21].

Future work. I plan to extend my existing work in this area in four ways. First, building upon initial positive feedback from law professors, I am developing a law-first version of this paper to better communicate the technical ideas to a non-technical audience, so as to increase the impact and the potential for adoption by courts. Second, I plan to write an amicus brief when this topic inevitably reaches the Supreme Court. My brief will argue for the consistency of our theory's approach with first principles of the Fifth Amendment. Third, our work in [CSV22][1] points to new opportunities for FC-resilient systems that I wish to unify with my existing work in this area [SV21, Scheffler21][1]. Finally, with the Supreme Court's ultimate decision in hand, we will need to interpret how that ruling should inform the design of current and future cryptosystems.

Other interdisciplinary works

I have also completed other works in the intersection of computer science and law. In [STV22][1], I built a formal framework for analyzing whether one work is derivative of another for the purposes of copyright law, based around the computational complexity idea of description length; I also presented this work at ACM CS/Law 2022. In 2019, my paper on autonomous weapon systems [SO19] won 2nd place at the inaugural ACM CS/Law Symposium's Student Paper Competition; I plan to update this work with some significant new changes to the ecosystem of autonomous weapons that have developed in the last two years. Around the same time I also conducted a study with a joint group of computer science and law professors to investigate how specific types of device fingerprinting were disclosed in privacy policies across the web [MSS$^+$21][1]. Finally, I wrote an algorithmic fairness paper on the difficulties moving from non-binary to binary classifiers [CCD$^+$19][1]; it was accepted at FAccT 2019, the premier algorithmic fairness conference.

2 Advances in Applied Cryptography

None of the methods described in the previous section would be possible without making advances in applied cryptography itself. My second research direction is to make foundational advances in traditional applied cryptography. Zero knowledge proofs [33, 13] allow a "prover" to convince a "verifier" of the truth of an NP statement, without revealing the witness that shows why the statement is true; for example, one might prove that a circuit is satisfiable without revealing the satisfying input. They are also key components of multi-party computation [32, 48], in which a collection of parties jointly compute the output to a function without revealing anything about their sensitive input data.

Both of these technologies are key components of the more socially beneficial uses of cryptography, Zero-knowledge proofs appear in use cases as far ranging as nuclear armament verification [31], crime scene DNA non-matches [26], and enforcing rules for data surveillance and warrants [46, 28, 46, 44]. Multi-party computation has a similarly impressive list of applications, including detecting tax fraud [14], privately computing prices in electricity markets [3], avoiding satellite collisions [41], and measuring the gender and racial pay gap in the city of Boston [11].

My research on these topics covers two areas: foundational improvements to zero-knowledge proofs, and novel applications of multi-party computation in other disciplines.

Concrete improvements to zero-knowledge proof efficiency. My first improvement to zero-knowledge proofs was BooLigero [GSV21][1], an adaptation of the Ligero proof system for arithmetic circuits of Ames et al. [5] to the Boolean setting which I presented at Financial Cryptography in 2021. For Boolean circuits like SHA-3, we achieved a reduction in proof size of $1.75 - 3\times$ over original Ligero without sacrificing prover or verifier runtime or memory. For hybrid arithmetic-Boolean operation circuits like SHA-2, we improved proof size by up to $1.6\times$.

I continued this line of research with a work accepted at Applied Cryptography and Network Security 2021 [GHS$^+$21][1]. Our new TurboIKOS system follows the "MPC-in-the-head" or "IKOS" framework of Ishai et al. [39], in which the prover commits to emulated executions of several MPC parties, and then opens some

of these to the verifier, who can use them to verify the prover's correct behavior with a probability based on the number of revealed parties. TurboIKOS remains state-of-the-art in MPC-in-the-head proofs: it improves upon Baum and Nof's prior work [10] incorporating the Beaver triple "sacrificing" approach of MPC for use in zero-knowledge proofs, achieving a proof size comparable to the "cut-and-choose" approach of Katz et al. [42] and the very different polynomial-interpolation approach of Baum et al. [9] but with significantly lower memory costs.

Multi-party computation for multi-agent robotics. In robotics, Simultaneous Localization And Mapping (SLAM) [51] is a task in which an agent builds a map of an unknown environment and tracks its position within it. When SLAM is extended to a multi-agent system [17], each agent builds its own local map and communicates with other agents to relate their positions in a shared reference frame. To compute this shared reference frame, the agents typically share a wealth of data (e.g. visual descriptors, odometry, or motion plans) in the clear, a needless loss of privacy for the individual agents. My ongoing work, planned for submission at Usenix Security 2023, brings cryptographic privacy to this process: agents use multi-party computation to perform this shared computation without revealing their extensive private data. Performing multi-agent SLAM in a privacy-preserving way is a sizeable effort in applied cryptography, utilizing not only standard efficient methods and libraries for multi-party computation [62, 25, 64] but also algorithms for secure ranking [47, 35], stochastic gradient descent [16, 49] and Oblivious RAM [6, 37, 34].

Future work. Under the topic of zero-knowledge proofs, the MPC-in-the-head paradigm features an intriguing parameter tradeoff: if the prover uses more emulated parties, the proof size will be lower, but the prover's runtime increases in the number of parties. I suspect creative use of homomorphism could combine the work done across multiple parties, yielding large efficiency gains.

For multi-party computation, techniques similar to the ones I have already worked with would apply well to other settings, bringing privacy to other kinds of navigation. This is a topic of paramount importance as the U.S. transitions toward a more cyber-physical world of self-driving cars, drones, and virtual reality. I am excited to be at the forefront of this area, and to use both my policy and cryptography expertise to navigate the challenges of the coming decade.

My references

[CCD+19] Ran Canetti, Aloni Cohen, Nishanth Dikkala, Govind Ramnarayan, Sarah Scheffler, and Adam Smith. From soft classifiers to hard decisions: How fair can we be? In *Proceedings of the conference on fairness, accountability, and transparency*, pages 309–318, 2019. Author order is alphabetical.

[CSV22] Aloni Cohen, Sarah Scheffler, and Mayank Varia. Can the government compel decryption? don't trust—verify. In *2nd ACM Symposium on Computer Science and Law*, 2022. Author order is alphabetical.

[GHS+21] Yaron Gvili, Julie Ha, Sarah Scheffler, Mayank Varia, Ziling Yang, and Xinyuan Zhang. Turboikos: Improved non-interactive zero knowledge and post-quantum signatures. In *International Conference on Applied Cryptography and Network Security*, pages 365–395. Springer, 2021. Author order is alphabetical.

[GSV21] Yaron Gvili, Sarah Scheffler, and Mayank Varia. Booligero: improved sublinear zero knowledge proofs for boolean circuits. In *International Conference on Financial Cryptography and Data Security*, pages 476–496. Springer, 2021. Author order is alphabetical.

[MSS+21] Julissa Milligan, Sarah Scheffler, Andrew Sellars, Trishita Tiwari, Ari Trachtenberg, and Mayank Varia. Case study: disclosure of indirect device fingerprinting in privacy policies. In *Socio-Technical Aspects in Security and Trust: 9th International Workshop, STAST 2019, Luxembourg City, Luxembourg, September 26, 2019, Revised Selected Papers 9*, pages 175–186. Springer, 2021. Author order is alphabetical.

[Scheffler21] Sarah Scheffler. password-ag2pc: An FC-reilient version of EMP-ag2pc., 12 2021. https://github.com/sarahscheffler/password-ag2pc.

[SKM23] Sarah Scheffler, Anunay Kulshrestha, and Jonathan Mayer. Public verification for private hash matching, 2023. Under submission at IEEE Security & Privacy.

[SM23] Sarah Scheffler and Jonathan Mayer. Systematization of knowledge: Content moderation for end-to-end encryption, 2023. Under submission at Privacy Enhancing Technologies Symposium.

[SMK22] Sarah Scheffler, Jonathan Mayer, and Anunay Kulshrestha. Comments on the safety tech challenge fund evaluation criteria, 4 2022. URL: https://sarahscheffler.net/Comments_on_UK_Safety_Tech_Challenge_Evaluation_Criteria.pdf.

[SO19] Sarah Scheffler and Jacob Ostling. Dismantling false assumptions about autonomous weapon systems. 2019.

[STV22] Sarah Scheffler, Eran Tromer, and Mayank Varia. Formalizing human ingenuity: A quantitative framework for coyright law's substantial similarity. In *2nd ACM Symposium on Computer Science and Law*, 2022. Author order is alphabetical.

[SV21] Sarah Scheffler and Mayank Varia. Protecting cryptography against compelled self-incrimination. In *30th USENIX Security Symposium (USENIX Security 21)*, 2021. Author order is alphabetical.

Other references

[1] Safety tech challenge fund. Technical report, United Kingdom, 2021. URL: https://www.gov.uk/government/news/government-funds-new-tech-in-the-fight-against-online-child-abuse.

[2] H. Abelson, R. Anderson, S. M. Bellovin, J. Benaloh, M. Blaze, J. Callas, W. Diffie, S. Landau, P. G. Neumann, R. L. Rivest, et al. Bugs in our pockets: The risks of client-side scanning. *arXiv preprint arXiv:2110.07450*, 2021.

[3] A. Abidin, A. Aly, S. Cleemput, and M. A. Mustafa. Towards a local electricity trading market based on secure multiparty computation. *COSIC internal report*, 2016.

[4] J. F. Almansa, I. Damgård, and J. B. Nielsen. Simplified threshold RSA with adaptive and proactive security. In *EUROCRYPT*, 2006. URL: https://doi.org/10.1007/11761679_35.

[5] S. Ames, C. Hazay, Y. Ishai, and M. Venkitasubramaniam. Ligero: Lightweight sublinear arguments without a trusted setup. In B. M. Thuraisingham, D. Evans, T. Malkin, and D. Xu, editors, *ACM CCS 2017: 24th Conference on Computer and Communications Security*, pages 2087–2104. ACM Press, Oct. / Nov. 2017.

[6] G. Asharov, I. Komargodski, W.-K. Lin, K. Nayak, E. Peserico, and E. Shi. Optorama: Optimal oblivious ram. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 403–432. Springer, 2020.

[7] F. Baldimtsi, J. Camenisch, M. Dubovitskaya, A. Lysyanskaya, L. Reyzin, K. Samelin, and S. Yakoubov. Accumulators with applications to anonymity-preserving revocation. In *IEEE European Symposium on Security and Privacy, EuroS&P*, 2017. URL: https://doi.org/10.1109/EuroSP.2017.13.

[8] F. Baldimtsi, R. Canetti, and S. Yakoubov. Universally composable accumulators. In *Topics in Cryptology - CT-RSA*. Springer, 2020. URL: https://doi.org/10.1007/978-3-030-40186-3_27.

[9] C. Baum, C. D. de Saint Guilhem, D. Kales, E. Orsini, P. Scholl, and G. Zaverucha. Banquet: Short and fast signatures from aes. In *Public Key Cryptography (1)*, pages 266–297, 2021.

[10] C. Baum and A. Nof. Concretely-efficient zero-knowledge arguments for arithmetic circuits and their application to lattice-based cryptography. In A. Kiayias, M. Kohlweiss, P. Wallden, and V. Zikas, editors, *PKC 2020: 23rd International Conference on Theory and Practice of Public Key Cryptography, Part I*, volume 12110 of *Lecture Notes in Computer Science*, pages 495–526. Springer, Heidelberg, May 2020.

[11] A. Bestavros, A. Lapets, and M. Varia. User-centric distributed solutions for privacy-preserving analytics. *Communications of the ACM*, 60(2):37–39, 2017.

[12] A. Bhowmick, D. Boneh, S. Myers, K. Talwar, and K. Tarbe. The apple psi system. Technical report, Apple, Inc., 2021. URL: `https://www.apple.com/child-safety/pdf/Apple_PSI_System_Security_Protocol_and_Analysis.pdf`.

[13] M. Blum, P. Feldman, and S. Micali. Non-interactive zero-knowledge and its applications (extended abstract). In *20th Annual ACM Symposium on Theory of Computing*, pages 103–112. ACM Press, May 1988.

[14] D. Bogdanov, M. Jõemets, S. Siim, and M. Vaht. How the estonian tax and customs board evaluated a tax fraud detection system based on secure multi-party computation. In *International conference on financial cryptography and data security*, pages 227–234. Springer, 2015.

[15] T. Brewster. Apple delays iphone child sexual abuse scanning after uproar. *Forbes*, 9 2021. URL: `https://www.forbes.com/sites/thomasbrewster/2021/09/03/apple-delays-iphone-child-sexual-abuse-scanning-after-uproar/?sh=4cfb0e4212d2`.

[16] N. Chandran, D. Gupta, A. Rastogi, R. Sharma, and S. Tripathi. Ezpc: programmable and efficient secure two-party computation for machine learning. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 496–511. IEEE, 2019.

[17] Y. Chang, Y. Tian, J. P. How, and L. Carlone. Kimera-multi: a system for distributed multi-robot metric-semantic simultaneous localization and mapping. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 11210–11218. IEEE, 2021.

[18] D. Chaum, J.-H. Evertse, and J. v. d. Graaf. An improved protocol for demonstrating possession of discrete logarithms and some generalizations. In *Workshop on the Theory and Application of of Cryptographic Techniques*, pages 127–141. Springer, 1987.

[19] R. Cohen. Asynchronous secure multiparty computation in constant time. In *PKC 2016*. Springer, 2016. URL: `https://doi.org/10.1007/978-3-662-49387-8_8`.

[20] E. Commission. Eu digital services act (dsa), 2022. URL: `https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en`.

[21] E. Commission. Regulation of the european parliament and of the council laying down rules to prevent and combat child sexual abuse, 2022. URL: `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A209%3AFIN&qid=1652451192472`.

[22] Commonwealth v. Davis, Pa: Supreme Court, Middle Dist. 2019.

[23] Commonwealth v. Gelfgatt, 468 Mass. 512, 11 N.E.3d 605, 11 N.E. (2014).

[24] Commonwealth v. Jones, 481 Mass. 540 - Mass: Supreme Judicial Court 2019.

[25] I. Damgård, V. Pastro, N. Smart, and S. Zakarias. Multiparty computation from somewhat homomorphic encryption. In *Annual Cryptology Conference*, pages 643–662. Springer, 2012.

[26] B. Fisch, D. Freund, and M. Naor. Physical zero-knowledge proofs of physical properties. In J. A. Garay and R. Gennaro, editors, *Advances in Cryptology – CRYPTO 2014*, pages 313–336, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

[27] Fisher v. United States, 425 U.S. 391, 96 S. Ct. 1569, 48 L. Ed. 2d 39 (1976).

[28] J. Frankle, S. Park, D. Shaar, S. Goldwasser, and D. J. Weitzner. Practical accountability of secret processes. In W. Enck and A. P. Felt, editors, *USENIX Security 2018: 27th USENIX Security Symposium*, pages 657–674. USENIX Association, Aug. 2018.

[29] T. K. Frederiksen, T. P. Jakobsen, J. B. Nielsen, and R. Trifiletti. On the complexity of additively homomorphic uc commitments. In *Theory of Cryptography Conference*, pages 542–565. Springer, 2016.

[30] D. K. Gillmor. Apple's new 'child safety' plan for iphones isn't so safe. *American Civil Liberties Union*, 8 2021. URL: `https://www.aclu.org/news/privacy-technology/apples-new-child-safety-plan-for-iphones-isnt-so-safe/`.

[31] A. Glaser, B. Barak, and R. J. Goldston. A zero-knowledge protocol for nuclear warhead verification. *Nature*, 510(7506):497–502, 2014.

[32] O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game or A completeness theorem for protocols with honest majority. In A. Aho, editor, *19th Annual ACM Symposium on Theory of Computing*, pages 218–229. ACM Press, May 1987.

[33] O. Goldreich, S. Micali, and A. Wigderson. Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. *Journal of the ACM*, 38(3):691–729, 1991.

[34] O. Goldreich and R. Ostrovsky. Software protection and simulation on oblivious rams. *Journal of the ACM (JACM)*, 43(3):431–473, 1996.

[35] M. T. Goodrich. Zig-zag sort: A simple deterministic data-oblivious sorting algorithm running in o (n log n) time. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 684–693, 2014.

[36] Graham, Blumenthal, Cramer, Feinstein, Hawley, Jones, Casey, Whitehouse, Durbin, Ernst, Kennedy, Cruz, and Grassley. Earn it act of 2020, 2020.

[37] P. Grubbs, A. Khandelwal, M.-S. Lacharité, L. Brown, L. Li, R. Agarwal, and T. Ristenpart. Pancake: Frequency smoothing for encrypted data stores. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 2451–2468, 2020.

[38] Ireland Department of Tourism, Culture, Arts, Gaeltacht, Sport and Media. Online safety and media regulation bill, 2022. URL: `https://www.gov.ie/en/publication/d8e4c-online-safety-and-media-regulation-bill`.

[39] Y. Ishai, E. Kushilevitz, R. Ostrovsky, and A. Sahai. Zero-knowledge from secure multiparty computation. In D. S. Johnson and U. Feige, editors, *39th Annual ACM Symposium on Theory of Computing*, pages 21–30. ACM Press, June 2007.

[40] S. Kamara, M. Knodel, E. Llansó, G. Nojeim, L. Qin, D. Thakur, and C. Vogus. *Outside Looking In*. Center for Democracy and Technology, 2021.

[41] L. Kamm and J. Willemson. Secure floating point arithmetic and private satellite collision analysis. *International Journal of Information Security*, 14(6):531–548, 2015.

[42] J. Katz, V. Kolesnikov, and X. Wang. Improved non-interactive zero knowledge with applications to post-quantum signatures. In D. Lie, M. Mannan, M. Backes, and X. Wang, editors, *ACM CCS 2018: 25th Conference on Computer and Communications Security*, pages 525–537. ACM Press, Oct. 2018.

[43] O. S. Kerr. Compelled decryption and the privilege against self-incrimination. *Tex. L. Rev.*, 97:767, 2018.

[44] J. A. Kroll, E. W. Felten, and D. Boneh. Secure protocols for accountable warrant execution, 2014. https://www.cs.princeton.edu/~felten/warrant-paper.pdf.

[45] A. Kulshrestha and J. Mayer. Identifying harmful media in end-to-end encrypted communication: Efficient private membership computation. In *30th USENIX Security Symposium (USENIX Security 21)*, 2021.

[46] A. Kulshrestha and J. Mayer. Estimating incidental collection in foreign intelligence surveillance:{Large-Scale} multiparty private set intersection with union and sum. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 1705–1722, 2022.

[47] W.-K. Lin, E. Shi, and T. Xie. Can we overcome the n log n barrier for oblivious sorting? In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2419–2438. SIAM, 2019.

[48] Y. Lindell and B. Pinkas. An efficient protocol for secure two-party computation in the presence of malicious adversaries. *Journal of Cryptology*, 28(2):312–350, Apr. 2015.

[49] P. Mohassel and Y. Zhang. Secureml: A system for scalable privacy-preserving machine learning. In *2017 IEEE symposium on security and privacy (SP)*, pages 19–38. IEEE, 2017.

[50] J. Mullin. In 2021, we told apple: Don't scan our phones. *Electronic Frontier Foundation*, 2021.

[51] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE transactions on robotics*, 31(5):1147–1163, 2015.

[52] B. H. of Commons. Online safety bill, 2021. URL: `https://bills.parliament.uk/bills/3137/publications`.

[53] M. O'Neill, S. Ruoti, K. Seamons, and D. Zappala. Tls inspection: how often and who cares? *IEEE Internet Computing*, 21(3):22–29, 2017.

[54] OpenNet Initiative. West censoring east: The use of western technologies by middle east censors. 2011. URL: `https://opennet.net/sites/opennet.net/files/ONI_WestCensoringEast.pdf`.

[55] T. P. Pedersen. Non-interactive and information-theoretic secure verifiable secret sharing. In *Annual international cryptology conference*, pages 129–140. Springer, 1991.

[56] R. Pfefferkorn. Content-oblivious trust and safety techniques: Results from a survey of online service providers. *Available at SSRN 3920031*, 2021.

[57] R. S. Raman, A. Stoll, J. Dalek, R. Ramesh, W. Scott, and R. Ensafi. Measuring the deployment of network censorship filters at global scale. In *NDSS*, 2020. URL: `https://www.ndss-symposium.org/wp-content/uploads/2020/02/23099-paper.pdf`.

[58] L. Sacharoff. Unlocking the Fifth Amendment: Passwords and encrypted devices. *Fordham Law Review*, 87:203–251, 2018.

[59] Seo v. State, 109 N.E.3d 418 (Ind. Ct. App. 2018).

[60] State v. Andrews, 197 A. 3d 200 - NJ: Appellate Div. 2018.

[61] X. Wang, A. J. Malozemoff, and J. Katz. EMP-toolkit: Efficient MultiParty computation toolkit, 2016. URL: `https://github.com/emp-toolkit`.

[62] X. Wang, S. Ranellucci, and J. Katz. Authenticated garbling and efficient maliciously secure two-party computation. In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security*, pages 21–37, 2017.

[63] A. T. Winkler. Password protection and self-incrimination: Applying the Fifth Amendment privilege in the technological era. *Rutgers Computer & Technology Law Journal*, 39:194–215, 2013.

[64] A. C.-C. Yao. How to generate and exchange secrets. In *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)*, pages 162–167. IEEE, 1986.