doi:10.1017/aju.2024.3

SYMPOSIUM ON DIGITAL EVIDENCE

THE ETHICAL AND LEGAL DILEMMAS OF DIGITAL ACCOUNTABILITY RESEARCH AND THE UTILITY OF INTERNATIONAL NORM-SETTING

Siena Anstis, Jakub Dalek,** and Ronald J. Deibert****

Nearly every aspect of our life is impacted by digital technologies manufactured and sold by companies. Legislative frameworks to limit the harms of such technologies have been slow to develop and remain entangled in controversy.[1] The expanding role of digital technologies has been accompanied by a disturbing descent into authoritarianism in many countries that is also, in part, fueled by these very same tools.[2] The decline of liberal democratic institutions is said to be linked to various properties of the digital ecosystem—from security flaws in popular applications used by states to engage in covert and remote surveillance[3] to the development and exploitation of social media algorithms that push violent and divisive content.[4] There is no doubt, then, that digital accountability research—which we define as evidence-based research seeking to track and expose risks to civil society in the digital ecosystem—is critical. This essay highlights the legal and ethical challenges faced in digital accountability research and concludes that a comprehensive and global ethical framework for such research is a critical step forward. As legal frameworks and norms continue to shift with respect to digital accountability research, such collaborative, international norm-setting would help ensure that digital accountability research continues.

Ethical and Legal Dilemmas: A Few Examples

Social scientists have widely observed the dual-edged implications of ongoing digital transformations. On the one hand, social media and other digital technologies vastly expand both the volume and types of data available to researchers, potentially leveling informational asymmetries in the process.[5] On the other hand, digital

* *Senior Legal Advisor, The Citizen Lab, Munk School of Global Affairs and Public Policy at the University of Toronto, Canada; PhD Fellow in Law, University of Oslo, Norway.*

** *Senior Network and Security Researcher, The Citizen Lab, Munk School of Global Affairs and Public Policy at the University of Toronto, Canada.*

*** *Director of the Citizen Lab and Professor of Political Science in the Munk School of Global Affairs & Public Policy at the University of Toronto, Canada.*

[1] Digital technologies here are defined widely to include anything from social media platforms to spyware companies. *See* RONALD DEIBERT, RESET: RECLAIMING THE INTERNET FOR CIVIL SOCIETY (2020).

[2] *See, e.g.*, Sarah Repucci & Amy Slipowitz, *Freedom in the World 2022: The Global Expansion of Authoritarian Rule*, FREEDOM HOUSE (2022). On the role of digital technology in authoritarianism, see, e.g., Marcus Michaelsen & Marlies Glasius, *Authoritarian Practices in the Digital Age: Introduction*, 12 INT'L. J. COMMC'N 3795 (2018).

[3] Ronald J. Deibert, *The Autocrat in Your iPhone: How Mercenary Spyware Threatens Democracy*, FOR. AFF. (Dec. 12, 2023).

[4] *Id.*

[5] Henning Lahmann, *Ukraine, Open-Source Investigations, and the Future of International Legal Discourse*, 116 AJIL 4 (2022).

accountability researchers (DARs)[6] must navigate the power disparities between themselves, impacted persons (for example, individuals whose phones have been hacked by a state actor), governments, and companies, and often in the absence of institutional guidelines or accountability mechanisms.[7]

DARs draw on a mix of ethical and legal frameworks in decision making around digital accountability research. Where digital accountability research involves interacting with individuals or has an impact on their privacy, these frameworks may include government guidelines on human subject research, which prioritize respect for persons, concern for welfare, and concern for justice.[8] International human rights law, such as the right to privacy and freedom of expression, may provide norms that govern the issues that DARs investigate or help to identify problematic uses of technology. Domestic law can help set limits on research techniques, but those legal frameworks may not adequately consider public interest research. Ethical frameworks for computer science research may also apply, but they may not cover important issues such as the tension between law enforcement investigations and responsible disclosure. Making matters more difficult, while digital technologies and their impact span jurisdictional boundaries, DARs may operate in isolation from other DARs, separated by geography or disciplinary boundaries. Finally, implementing general principles and frameworks is in practice complex.

One cutting-edge area of digital accountability research that illustrates a myriad of ethical challenges is tracking the commercial surveillance marketplace and documenting government digital espionage that targets civil society using mercenary spyware tools.[9] The very act of publishing research on this topic is inherently disruptive: it exposes secret government espionage campaigns and the products, tools, and services that are supplied to state actors by surveillance firms to help carry out these campaigns. It alerts unsuspecting individuals and organizations that they may be under surveillance, which can lead to calls for regulatory responses or legal action.[10]

DARs must often balance their dual concerns for the public's digital security and the use of such technologies by law enforcement and intelligence agencies. For example, researchers have acquired copies of sophisticated methods employed by spyware firms to hack into widely used software platforms.[11] There is an ongoing effort by DARs (including the authors' research group) to develop a standard practice in spyware research for DARs' responsible disclosure of technical details to the software vendor.[12] Responsible disclosure[13] generally leads to the patching of vulnerabilities (including zero-days[14]) by the software vendor, improving digital security for all. However,

[6] We define Digital Accountability Researchers (DARs) as researchers who undertake evidence-based research seeking to track and expose risks to civil society (defined broadly to include human rights defenders, dissidents, activists, and journalists) in the digital ecosystem. DARs include academics and members of civil society who are independent of both governments and private corporations.

[7] Zara Rahman & Gabriela Ivens, *Ethics in Open Source Investigations, in* DIGITAL WITNESS: USING OPEN SOURCE INFORMATION FOR HUMAN RIGHTS INVESTIGATION, DOCUMENTATION AND ACCOUNTABILITY (Sam Dubberley, Alexa Koenig & Daragh Murray eds., 2019).

[8] Government of Canada, TCPS 2 (2022): Chapter 1: Ethics Framework (2022). In the U.S. context, see U.S. Department of State Health and Human Services, The Belmont Report (1979).

[9] See Deibert, *supra* note 3. Mercenary spyware tools encompass spyware technology that is made and sold to government actors by the private sector and is used by law enforcement, military, and intelligence bodies.

[10] *See, e.g.*, Bill Marczak, John Scott-Railton, Adam Senft, Bahr Abdul Razzak & Ron Deibert, *The Kingdom Came to Canada: How Saudi-Linked Digital Espionage Reached Canadian Soil*, CITIZEN LAB (Oct. 1 2018).

[11] *See, e.g.*, Bill Marczak & John Scott-Railton, *The Million Dollar Dissident: NSO Group's iPhone Zero-Days Used Against a UAE Human Rights Defender*, CITIZEN LAB (Aug. 24, 2016).

[12] *Vulnerability Policy*, CITIZEN LAB (Nov. 19, 2020).

[13] Jonathan W. Penney, *Code Is Law, But Law Is Increasingly Determining the Ethics of Code*, MEDIUM (Jan. 22, 2015).

[14] U.S. Department of Commerce National Institute of Standards and Technology, Computer Security Resource Center (CSRC), Zero Day Attack (n.d.). A "zero day" refers to a software vulnerability or security flaw that is exploited by attackers before a patch is available or the software vendor is aware of the vulnerability, giving users zero days to defend against it.

https://doi.org/10.1017/aju.2024.3 Published online by Cambridge University Press

government agencies may be using these same vulnerabilities and responsible disclosure could impact an ongoing investigation. Similarly, DARs notifying human subjects that their devices have been targeted or infected with mercenary spyware can be invaluable for them—but such notifications could also disrupt a law enforcement or intelligence investigation.

Applying the guiding principles of respect for persons and concern for justice can help navigate DARs' decision making with reference to broader principles that go beyond the specific case at issue. Application of this ethical framework values the public's digital security over law enforcement and intelligence activities if such targeting would likely violate international human rights law (for example, the targeting of a journalist or human rights defender with spyware and where there is no reasonable basis to conclude that such targeting is prescribed by law, legitimate, necessary, and proportionate).[15] Such an ethical approach is in part further justified because state agencies, whether democratic or not, are subject to little oversight and transparency in their use of spyware. Without notification that their device has been targeted or infected with spyware, the targeted person would never know because states do not generally disclose such information to the victim.[16]

Peer Review

Publication of technical findings related to spyware investigations also raises further ethical challenges. While transparency is critical to academic work, revealing technical findings may also contribute to the continuation of harm by giving mercenary spyware companies the information they require to further adapt their systems and evade research. Due to these concerns, DARs may opt not to publish certain pieces of evidence that would compromise ongoing research. This decision making prioritizes the possibility of further justice for targets of spyware if additional avenues of targeting can be detected by keeping technical details from spyware vendors and ensuring that additional research avenues remain open. To honor the principle of academic transparency, while also addressing these risks, DARs may engage in alternative forms of peer review such as sharing details with other experts using the Traffic Light Protocol (TLP) developed by the Forum of Incident Response and Security Teams (FIRST) for the sharing and critical evaluation of sensitive information.[17] For example, technical information that is shared with a peer reviewer and designated TLP:RED should only be viewed by that specific individual and should not be shared with anyone else.

Limits Around Network Interrogation

A third area where DARs encounter ethical issues concerns the identification of networking equipment on the networks of internet service providers that may be implicated in censorship or surveillance. When undertaking this work, DARs will often interact with publicly accessible computer systems which form a part of the infrastructure related to censorship and surveillance. Such interaction can include accessing block page servers and administration and analytics interfaces of networking devices. These devices are often actively being used on the internet service providers' networks. For example the same devices that can block access to websites can often be used for caching content to improve network speed. Ethical considerations require that DARs proceed cautiously when interacting with any computer system that they do not own and avoid negatively affecting these systems or the

[15] Human rights groups have also argued that mercenary spyware is inherently at odds with international human rights law. *See* European Data Protection Supervisor, Preliminary Remarks on Modern Spyware (2022).

[16] *See, e.g.*, *NSO Group/Q Cyber Technologies: Over One Hundred New Abuse Cases*, Citizen Lab (Oct. 29, 2019).

[17] CISA, *Traffic Light Protocol 2.0 User Guide* (2022).

https://doi.org/10.1017/aju.2024.3 Published online by Cambridge University Press

experience of other users.[18] Ethical researchers will also often follow the injunction to not attempt to circumvent any security measures on a device (and, indeed, this may be a legal requirement depending on the jurisdiction).

Institutional and Legal Support Around Digital Accountability Research

DARs need appropriate institutional and legal support to address these ethical challenges. This section highlights some of the key elements of the institutional architecture needed for digital accountability research, with a particular focus on DARs in the university context. With minor adjustments, these institutional frameworks can be adapted to other research contexts.

Research ethics protocols, an institutionalized approach to research in universities, can help guide the collection of personal information and its secure holding. These protocols ensure that human research subjects who are affected by digital accountability research are able to provide informed consent to the risks that they may face in being the subject of such research. Further, research ethics protocols are reviewed and approved by a university-level Research Ethics Board, which provides an important independent check on university-based DARs. There are compelling reasons for researchers to ensure compliance with such protocols even in situations that do not necessarily require it (such as pure technical research or research outside the university context) in order to think through privacy and security issues.[19] At the core of the research ethics procedure is the need for researchers to define the scope of their research, justify what they are doing, explain their approach, and consider any potential harm to persons and how it can be mitigated. These processes help anticipate and reduce potential harm to specific persons, including those that may arise inadvertently because of poorly secured technical data that may contain personally identifiable information.[20]

Conducting ethical research also requires robust institutional and community support.[21] Legal counsel need to be socialized into the ethical and legal dilemmas that are raised by digital accountability research and not be afraid to take a bold approach to defending research that is in the public interest. Organizations like the Electronic Frontier Foundation or others providing *pro bono* legal support to digital accountability researchers play a critical role in situations where litigation by companies can decimate DARs' work. More broadly, there is a role for global community networks in confronting challenges faced by DARs. As we discuss in the conclusion, more efforts could be undertaken to bring together the research community and to debate and build internal norms on digital accountability research.

There are ethical and legal issues around digital accountability research that relate not just to subjects of the research but to the researchers themselves. For example, defamation suits are among the common tactics employed by companies to silence DARs.[22] However, the growth of anti-strategic litigation against public participation (anti-SLAPP) laws has made it more difficult for companies to pursue and shut down public interest research.[23] Laws intended to protect the public, however, may also, at times and depending on who is interpreting and applying them,

[18] Eric Pauley & Patrick McDaniel, *Understanding the Ethical Frameworks of Internet Measurement Studies*, IEEE SECURITY & PRIVACY (2017).

[19] *See, e.g.*, Anne E. Boustead & Trey Herr, *Analyzing the Ethical Implications of Research Using Leaked Data*, 53 POL. SCI. & POL'Y 3 (2020); Marcello Ienca & Effy Vayena, *Ethical Requirements for Responsible Research with Hacked Data*, 3 NATURE MACHINE INTELLIGENCE 9 (2021).

[20] Rosanna Bellini et al., *SoK: Safer Digital-Safety Research Involving At-Risk Users*, 13 (IEEE Symposium on Security and Privacy (Pre-Print)).

[21] *Id.*

[22] *See, e.g.*, Complaint, *X Corp v. Center for Countering Digital Hate, Inc., et al.* (N.D. Cal., July 31, 2023); Penney, *supra* note 13; Vittoria Elliott, *How X Is Suing Its Way Out of Accountability*, WIRED (Aug. 15, 2023). At the core of the litigation is the use of website scraping as a data collection method, which is currently prohibited by X's Terms of Service.

[23] *See* Centre for Free Expression, *Anti-SLAPP Legislation: A Backgrounder*, CFE (Mar. 19, 2019).

https://doi.org/10.1017/aju.2024.3 Published online by Cambridge University Press

unhelpfully hinder public interest research and should be challenged.[24] In particular, some laws pose a substantial risk when there are no specific, legislated protections for DARs.[25] The Computer Fraud and Abuse Act in the United States, for example, has been criticized for providing a basis for criminalizing digital accountability research in the absence of protections for good-faith security researchers.

One problematic example concerns the Canadian Panel on Research Ethics (which provides the ethical framework for university-level research in Canada). The panel issued guidelines that researchers should not violate the terms of service of platforms on which they are conducting research.[26] This guidance serves to limit research and helps social media platforms to insulate their services from public scrutiny. Specifically, automated website scraping is often a technical breach of the terms of service, done in the pursuit of public interest research and it is one of the few methods available to researchers to ensure platform accountability. Terms of Service that prohibit this practice should not be allowed to limit platform accountability. This case serves as another reminder that researchers need to proactively shape the social, legal, and ethical direction of their fields before legal regimes harden in ways that may limit the scope of such legitimate inquiries.[27]

Concluding Observations: The Need for International Norm-Setting

Digital accountability research is of increasing importance and researchers are facing more complex challenges. There is a critical role for a broader cross-jurisdictional debate on these issues and the development of global norms.[28] The Berkeley Protocol on Digital Open-Source Investigations, which provides a set of principles and guidance for researchers to consider, is an interesting example of the potential utility of international norm-setting in research through a collaborative process supported by an international body.[29]

As a first step toward such global norm-setting for digital accountability research, we suggest that the Protocol could be taken as a template for the development of a broader set of international norms around digital accountability research that go beyond open-source investigations. Such a document could address spyware investigations, network censorship measurement, and disinformation and misinformation research. The deliberative process to generate these norms would also allow DARs to exchange views on ethical decision making and legal research constraints and develop cross-border strategies for pushing back against laws and other norms that constrain such research despite the clear public interest element. A core weakness is that researchers continue to operate largely in isolation in their respective jurisdictions (or even within their own institution). A global framework could also inform how various bodies, such as research ethics boards, universities, and governments, evaluate how researchers address ethical dilemmas and approach digital accountability research.

[24] Ewen Macaskill & Gabriel Dance, *NSA Files: Decoded*, GUARDIAN (Nov. 1, 2013).

[25] *See* Andrew Crocker, *Scraping Public Websites (Still) Isn't a Crime, Court of Appeals Declares*, ELECTRONIC FRONTIER FOUNDATION (Apr. 19, 2022).

[26] Government of Canada, Panel on Research Ethics, *Does Research Using Social Media Platforms Require Research Ethics Board Review?* (2022).

[27] Penney, *supra* note 13.

[28] Bellini et al. identify a number of safety practices that could inform such a framework and argue that more work is required to cultivate best practices. *See* Bellini et al., *supra* note 20, at 13.

[29] BERKELEY PROTOCOL ON DIGITAL OPEN-SOURCE INVESTIGATIONS: A PRACTICAL GUIDE ON THE EFFECTIVE USE OF DIGITAL OPEN-SOURCE INFORMATION IN INVESTIGATING VIOLATIONS OF INTERNATIONAL CRIMINAL, HUMAN RIGHTS AND HUMANITARIAN LAW (United Nations & University of California, Berkeley eds., 2022).

https://doi.org/10.1017/aju.2024.3 Published online by Cambridge University Press