Selected Papers of #AoIR2024:
The 25th Annual Conference of the
Association of Internet Researchers
Sheffield, UK 30 Oct - 2 Nov 2024

HACKERS' PRIVACY APPROCHES: HOW PRIVACY VIOLATION AND PRIVACY PROTECTION GO HAND IN HAND

Keren Levi-Eshkol
University of Haifa

Rivka Ribak
University of Haifa

Our study focuses on the enigmatic world of hackers and the seeming paradox of their engagement in both invading and safeguarding privacy. Using qualitative content analysis and code analysis of GitHub code projects (repositories), we seek to understand the cultural logic underlying this duality. Drawing on digital materialism, the study aims to shed light on the ethical implications and material significance of code, emphasizing the ongoing need to consider digital privacy ethics and practice in the ever-evolving technological landscape.

Over the last few decades, hacking has become a significant public concern. Hackers are involved in cyber incidents that result in substantial technological and financial repercussions for companies, organizations, governments, and individuals. Generally, a hacker is characterized as a technologist with a strong affinity for computing, programming, system/network administration, and security, often expressing their wit and humor through source code. A hack is recognized as a clever technical solution devised through unconventional and non-obvious means, combining craftsmanship with craftiness (Levy, 1984; Coleman, 2015; Wagen, 2018; Feiten & Coleman, 2022). The hackers' community transformed significantly from an underground subculture in the 1980s to today's acclaimed and professionalized domain. During this period, the emergence of imaginary hats became a symbolic representation of hacker roles, distinguishing "black hat" (associated with malicious hacking); "white hat" (hackers that are more cognizant and respectful of rules and regulations); and "gray hat" (hackers who may historically or presently utilize methods, such as breaking into systems, to comprehend and enhance security practices) (Wagen, 2018; Feiten & Coleman, 2022). Despite the variety of practices and ethics within the community, the unifying thread is

Suggested Citation (APA): Levi-Eshkol, K., Ribak, R. (2024, October). Hackers' privacy approaches: How privacy violation and privacy protection go hand in hand. Paper presented at AoIR2024: The 25th Annual Conference of the Association of Internet Researchers. Sheffield, UK: AoIR. Retrieved from http://spir.aoir.org.

the commitment to a free flow of information and the belief that computers can contribute to creating a better world.

Hackers often argue that they adhere to an ethic that justifies their unauthorized access to systems. According to Steinmetz and Gerber (2014; see also Levy, 1984), hacker ethics involve a mixture of two visions: a liberal-oriented advocacy for the belief that information should be freely accessible, and a technology-oriented desire to be hands-on with systems of all sorts in order to interact and understand technology better. A common argument among hackers is that those who breach systems are performing a service by exposing security flaws, and therefore, they should be encouraged or even rewarded (Spafford, 2017). This study examines these dual commitments and seeks to explore how hackers negotiate privacy and its violation in their code projects.

The study draws on digital materialism as a theoretical framework for understanding the evolving relationship between code, materiality, and culture. Digital materialism suggests that code has a material presence and cultural significance beyond its functional role, and should be viewed as a material artifact with inherent qualities (Fuller, 2008). To study code, we chose the open-source platform GitHub.com as our corpus because of its popularity among hackers (Coleman, 2013; Wagen, 2018). GitHub.com is a leading site for open-source development, providing developers with features for storing, sharing, and working on collaborative projects. Established in 2008 and acquired by Microsoft in 2018, GitHub.com reported hosting over 420 million repositories in 2023, created by a community exceeding 100 million developers globally, with a predominant presence in the US, India, and China.[1]

To explore the hackers' privacy negotiations, we searched for repositories with 'Stealer' in their name or description. 'Stealer' is a common term for describing a malicious code whose only purpose is to hack into other people's computers and steal their private information (e.g. *"An example evil app that tries to steal personal data…"; "An Android malware stealing user privacy…"; "A Chrome extension that will steal literally everything it can…"*). Our search yielded a compilation of 2,500 stealer repositories along with their respective owners. Subsequently, we searched for repositories that belong to the same hacker profile, that contain the term 'privacy'. We ensured that these repositories were not of malicious intent and that they were not copied ('forked') from other developers (a common practice in open-source platforms). The result was a much-reduced list of 52 hackers who crafted both malicious 'stealer' repositories, and non-malicious repositories across diverse topics that included privacy protection methods. We then conducted qualitative discourse and code analyses, which allowed us to study unobtrusively how hackers conceptualize and implement privacy.

To understand how hackers rationalize publishing malicious code that may risk privacy, we analyzed how they explained it to others, using discourse analytic tools to study the repositories' descriptions. We found that the hackers describe the malicious code as an exploration born out of curiosity and commitment to uncovering security vulnerabilities. They justify their projects as educational tools, often cautioning against illegal use – *"This is for educational purposes only, to understand how malware works. Do not use it*

[1] https://kinsta.com/blog/github-statistics/

maliciously or on any machines that you do not have permission for." To understand the hackers' privacy approaches, we analyzed what methods they use to protect privacy in the non-malicious code projects. We found that the hackers' logic that those who possess private information are responsible for its protection (Steinmetz & Gerber, 2014; Coleman, 2015; Horstmann, 2022) is materialized in code using two privacy approaches – privacy-by-policy and privacy-by-design (Spiekermann & Cranor, 2009). The perception that the end-user is responsible for privacy is translated to code that implements privacy-policy consent forms and permission-based access settings, such as those found in most social networks. The end-users are granted the authority to consent to or dissent from the use of their information or limit who may have access to it. However, once information departs from the end-user, the hackers relegate the responsibility for protecting privacy to the entity that owns the system that collects the data. Our findings suggest that this perception is translated to code by employing sophisticated protection methods such as homomorphic encryption and the decentralization of private information. These may indicate the hackers' perception that, as the creators of the software, they share the responsibility for protecting the end-users' privacy.

Surprisingly, most hackers have chosen to disclose their details on their GitHub profiles, sharing information such as their name, photo, email, and updated resume. One possible explanation is that they do not perceive themselves as malicious actors ("black hat" hackers) but rather as experts in privacy and security risks.

The study seeks to understand how hackers translate their liberal ethics and perspective on privacy dynamics into the code they develop. Hackers adhere to the belief that the responsibility for privacy lies with the owner of the information, be it the end-user or the software owner managing the information post-collection. This logic manifests in the absence of a perceived contradiction: hackers do not see themselves as violating privacy when crafting code that breaches it. Instead, they view the owner of the information as accountable, since they did not adequately secure the data, while the hacker merely revealed a potential privacy risk. Their code mirrored this logic by incorporating protective features that empower information owners to preserve users' privacy.

References

Coleman, G. (2015). *Hacker, hoaxer, whistleblower, spy: The many faces of Anonymous.* Verso books.

Feiten G., Coleman G. (2022). A dive into the hacking sea (interview with Pr. Gabriella Coleman). *Journal of Openness, Commons & Organizing*, https://hal.science/hal-03831387

Fuller, M. (2008). *Software studies: A lexicon*, Cambridge: MIT Press.

Horstmann, N. D. (2022). The power to selectively reveal oneself: Privacy protection among hacker-activists. *Ethnos*, *87*(2), 257-274.

Levy, S. (1984). *Hackers: Heroes of the computer revolution*. Garden City, NY: Anchor Press/Doubleday.

Spafford, E. H. (2017). Are computer hacker break-ins ethical? In J. Weckert, ed., *Computer ethics* (pp. 293-299). Routledge.

Spiekermann, S., & Cranor, L. F. (2009). Engineering privacy. *IEEE Transactions on software engineering, 35*(1), 67-82.

Steinmetz, K., & Gerber, J. (2014). "It doesn't have to be this way": Hacker perspectives on privacy. *Social Justice, 41*(3), 29-51.

van der Wagen, W. (2018). The Cyborgian deviant: An assessment of the hacker through Actor-Network Theory. *Journal of Qualitative Criminal Justice and Criminology, 6*(2), 157-178.