IDENTIFYING WITH PRIVACY:
REFERENCES TO PRIVACY IN DEVELOPERS' GITHUB PROFILES

Keren Levy-Eshkol
Department of Communication, University of Haifa

Rivka Ribak
Department of Communication, University of Haifa

In a competitive environment in which one of the marks of good developers is their proficiency in the extraction and analysis of users' personal information, how do developers negotiate the cultural value of privacy? This study, now underway, is part of a larger project that investigates "privacy" closest to where it is materialized in code – namely, in the open source code-sharing platform GitHub. GitHub opened in 2008 and has been owned by Microsoft since 2018. According to recent data, it is hosting more than 62 million users[1] and 200 million code repositories (36 million of which are public[2]) – the most widely used platform for open-source software development and more recently, talent recruitment. Using GitHub REST API and Python scripts written for this purpose, we extracted from developers' profiles 2025 bio descriptions containing the word "privacy." These form the corpus of the current analysis, in which we map articulations of privacy in developers' self-presentation to study how they negotiate the potential tension between their commitment to privacy and their professional employability.

Mapping privacy in developers' self-presentation

Conceptualizing developers as cultural producers – mediators who translate cultural values into the code that in turn becomes culture (Ribak, 2019) – GitHub provides us with access to these producers unobtrusively, in their "natural" environment. GitHub is a hub of code development, hosting developers who advocate open-source software, commercial enterprises employing open-source solutions, and talent scouts seeking potential additions to their teams. As such, it plays an important role in shaping

---

[1] https://github.com/search?q=type:user&type=Users
[2] https://venturebeat.com/2018/11/08/github-passes-100-million-repositories/; https://github.com/search?q=is:public

developers' personal and professional identities, while allowing us to explore how developers curate their online personas and manage the two (Flisfeder, 2015).

To analyze developers' usage of the term "privacy" within their bios, we used GitHub search tools and developed custom Python scripts to identify profile descriptions containing references to "privacy" in the text. We collected 2025 profiles generated between 2008 and 2021 that contained the word "privacy," and saved them in a single file that formed the dataset for the automated analysis. Using VOSviewer software, an NLP tool designed for constructing and visualizing bibliometric networks (van Eck & Waltman, 2010), we identified four distinct clusters that, based on the words in each cluster, we labeled: "freedom" (in red), "privacy advocacy" (green), "academia" (blue), and "professional experience" (yellow).

Figure 1: Developers' profile words associated with the word "privacy"
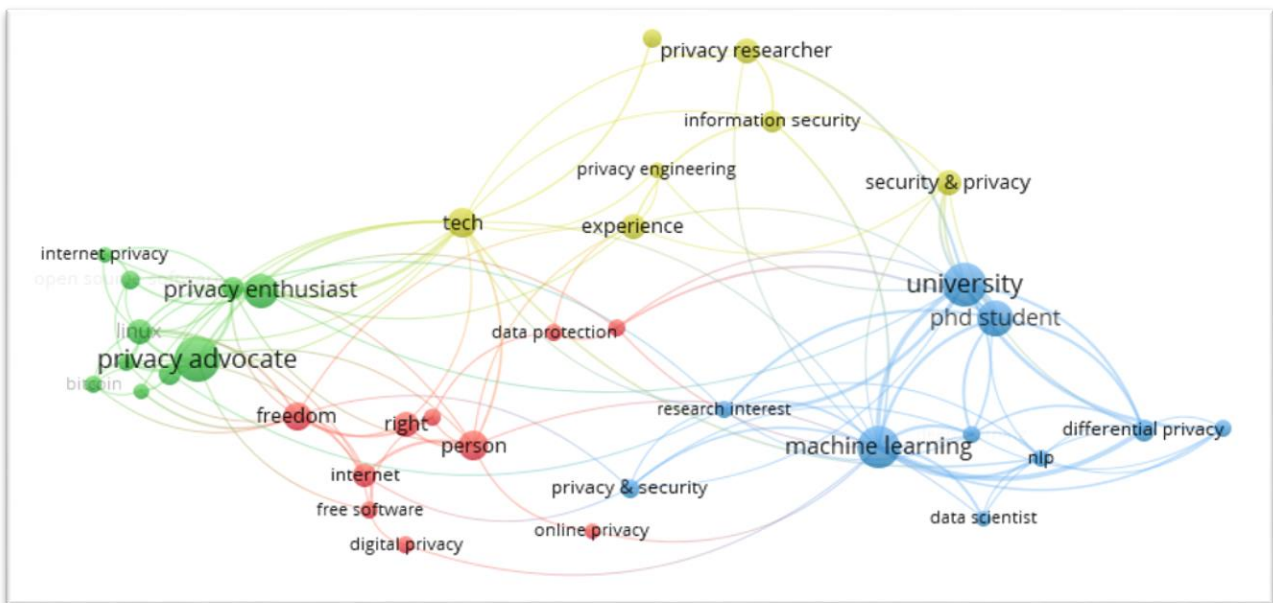

The proximity of the red/freedom and green/advocacy clusters suggests that developers who prioritize privacy as a human right tend to intertwine it with open-source practices. By contrast, the interconnections between the blue/academia and yellow/professional clusters imply that certain developers include privacy as part of their learned skills.

The Privacy Expert as a Freedom Fighter
To study these relationships further, we analyzed the interconnections of "freedom" and "experience" and their relationships with words from other clusters.

By pinpointing "freedom" on the VOSviewer map, we discerned its associations with various words and word pairs. Notably, "freedom" exhibited the strongest co-occurrence with terms and word pairs found in the green/advocacy cluster, such as "hacker," "Linux," "FOSS," "bitcoin," and "decentralization." Additionally, it was associated with words from the red/freedom cluster, including "right," "human right," "internet," and "free software," as illustrated in Figure 2.

Figure 2: Links associated with the word "Freedom"
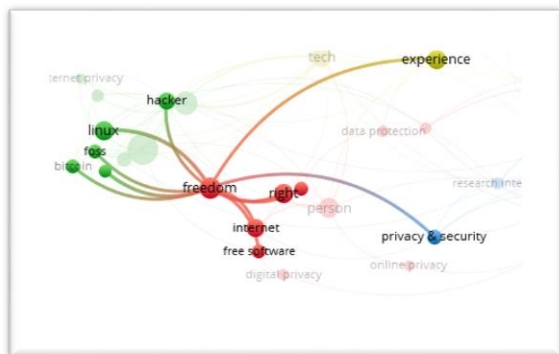
These relationships between "freedom" and green/advocacy cluster suggest that developers link open source to freedom, particularly in the context of privacy. Using voyant-tools.org, a scholarly text-analysis software that provides a Key Word In Context (KWIC) interface for interaction with the text, we identified profile descriptions that explicitly articulate this connection and include words from both clusters: "*Freedom and privacy geek - Open Source the world!*"; "*Security Engineer and Software Developer Enthusiast. … Believe in Privacy, Open Source, and Freedom.*" Furthermore, these profiles claim that open source code fosters privacy, in contrast to proprietary software, which is often perceived as a "black box": "*I am an avid Linux user. I believe that users have the right to privacy and a way to use their system the way they see fit and not the way a company tells you how.*"

When we focus on the term "experience" on the VOSviewer map, we find associations with words and word pairs such as "Ph.D. student," "privacy enthusiast," and "'freedom."

Figure 3: Words associated with the word "Experience"
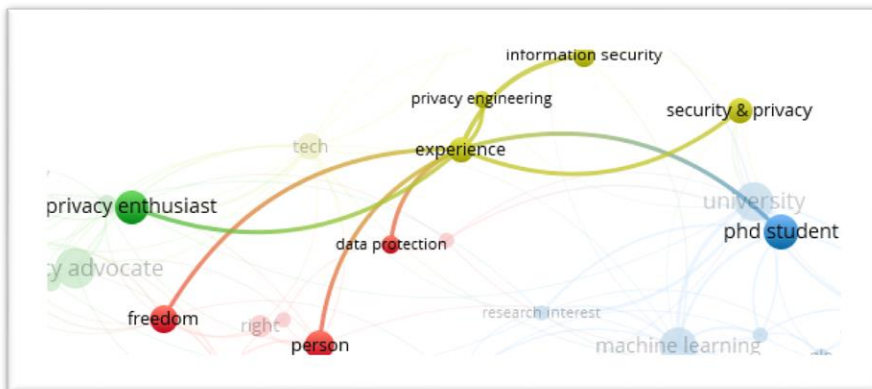
Beyond highlighting their professional skills, some developers also passionately express their dedication to privacy: "*C/C++ and Python dev. Advanced Linux/Unix experience. Privacy enthusiast.*" The use of terms like "freedom" and "privacy enthusiast" is not trivial, considering the extraction imperative (Zuboff, 2022) that undergirds the digital economy. Moreover, in contrast to the conventional linkage between employability and

a business-oriented mindset (Gershon, 2017), these developers do not perceive a conflict between the requirements imposed by software companies to align with their business models, and their belief in the value of privacy and its role in preserving freedom. Rather than focus exclusively on their professional accomplishments and skills, these developers opt for what may be termed authentic self-presentation, emphasizing privacy and the will to safeguard it: *"Graphic designer and web developer. Supporter of privacy, civil rights, and online liberties. Currently working at ..."* – arguably seeking employment in a workplace that values the privacy of its users (Sieverse et al., 2015).

"*Privacy is a human right. Encrypt everything*"

Finally, in three out of four clusters, the term "security" appears in different forms, including "information security," "privacy & security," "security & privacy," and "data protection." This finding echoes prior research that indicated that developers frequently employ security tools, notably encryption, for safeguarding privacy (Hadar et al., 2018). Consequently, their perception of privacy may primarily center on security concerns. As one developer wryly observed: "*Privacy is a human right. Encrypt everything.*"

We conclude by noting the need for a closer reading of the profiles, as the following attempts to reconcile the committed and the professional positions illustrate: *"By day, staff engineer; by night, privacy advocate,"* and *"Software engineer, privacy first."* The day/night distinction offers a temporal means to contain professional demands and personal activism within one person. Alternatively, by prioritizing privacy, the second bio outlines a hierarchy in which passion informs one's professional choices.

Cited References

Flisfeder, M. (2015). The entrepreneurial subject and the objectivization of the self in social media. *South Atlantic Quarterly*, *114*(3), 553-570.

Gershon, I. (2017). *Down and out in the new economy: How people find (or don't find) work today*. University of Chicago Press.

Hadar, I., et al. (2018). Privacy by designers: Software developers' privacy mindset. *Empirical Software Engineering*, *23*(1), 259-289.

Ribak, R. (2019). Translating privacy: Developer cultures in the global world of practice. *Information, Communication & Society*, *22*(6), 838-853.

Sievers, K., et al. (2015). Self-presentation in professional networks: More than just window dressing. *Computers in Human Behavior*, *50*, 25-30.

van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *scientometrics*, *84*(2), 523-538.

Zuboff, S. (2022). Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization. *Organization Theory*, *3*(3).