Whose Policy? Privacy Challenges of Decentralized Platforms

SOHYEON HWANG*, Northwestern University, USA
PRIYANKA NANAYAKKARA, Northwestern University, USA
YAN SHVARTZSHNAIDER, York University, Canada

In this position paper, we examine 100 privacy policies on Mastodon, a decentralized social networking platform, to reflect on the relationship between policy and technology design. The resurgence of interest in decentralized platforms is closely tied to concerns of data governance and user autonomy on their large, centralized counterparts. A key advantage of decentralized platforms is the opportunity for users to tailor the privacy policies and protocols governing their information and content. Despite this, servers in our sample overwhelmingly use the same generic privacy policy text. Furthermore, ambiguities in the dominant text prompt questions about whether implementations of the same policy differ in practice, how generic policies may fail to reflect user preferences, and what challenges arise when coordinating a decentralized network of communities and their policies.

1 THE RESURGENCE OF DECENTRALIZED SOCIAL NETWORK PLATFORMS

The ongoing ethical and regulatory challenges surrounding popular social network platforms have rekindled interest in decentralized alternatives [2]. While decentralized platforms may resemble their centralized counterparts, they embody significantly different organizing and governing principles. Large, centralized platforms like Twitter and Facebook are owned by privately-owned companies, run on proprietary servers and code, and centrally determine and enforce a broad set of rules and terms through design choices, algorithms, and moderation. In contrast, decentralized social networks like Mastodon rely on open-source software to form a federated network of self-hosted servers (i.e., "instances") [5]. Owners of servers are free to set moderation rules, privacy policies, and encoded controls reflecting their personal or shared community values (e.g., details around how data are retained) [1]. The late 2022 exodus from Twitter to Mastodon [see 3] brings attention to the question of how and whether communities leverage this greater degree of autonomy on decentralized networks to govern their online experiences by tailoring their respective rules and protocols.

2 CASE STUDY OF PRIVACY POLICIES ON MASTODON

We look to the privacy policies on Mastodon as a case study of how norms and institutions are technologically mediated across organically emerging, decentralized communities. To this end, we examined 100 English-language privacy policies from Mastodon servers randomly sampled from `instances.social`.[1] Our initial results show that, despite the ability to customize privacy policy, most server owners use generic privacy policy text without further adaptation. Removing server names from policy text resulted in just 11 unique texts: 81% of servers used the exact same text, with an additional 3% using a minor variation; another 12% of servers used highly similar variations of another text.

The homogeneity of privacy policies might suggest that users' privacy expectations are well-aligned across servers. However, the generic texts are often broad and ambiguous, leaving open many potential interpretations. It is unclear whether privacy *practices* differ for communities using the same text and whether the dominant policy text appropriately reflects privacy expectations on Mastodon. The federated design of Mastodon further contributes to this ambiguity. By design, Mastodon's network of servers form many distinct and overlapping (social) contexts a post can traverse. An ambiguous policy (see excerpt from the prevailing privacy policy below) about a server's relationship to "other servers" could lead to inadvertently misleading system behaviour, e.g., sharing posts with an unintended audience:

*Corresponding author: sohyeonhwang@u.northwestern.edu
[1] https://instances.social/api/doc/

> In some cases [posts] are delivered to different servers and copies are stored there. We make a good faith effort to limit the access to those posts only to authorized persons, but other servers may fail to do so. [...] Do not share any sensitive information over Mastodon.

This illustrates a critical privacy concern in maintaining the *contextual integrity* [4] of online content: while privacy practices can be tailored to a server, the multitude of servers means that data could be subjected to many policies, resulting in inevitable ambiguity in the governance of any given post.[2] As privacy is highly contextual [4], a user's level of comfort with sharing a specific piece of content may differ based on server's location, owner, and community norms.

3 UNDERSTANDING PRACTICES ACROSS NETWORKS OF POLICIES

We invite the research community to consider how technologies shape particular arrangements of policy across communities and vice versa through three potential research directions.

 # 1: Understanding current norms. First, while current privacy policies on Mastodon tend to be very similar, they are sufficiently ambiguous in ways that allow for varying implementations and practices. How and when do these ambiguities violate privacy expectations? To begin answering this question, we propose leveraging the framework of contextual integrity [4] to assess the appropriateness of information flows by specifying key parameters defining and shaping said flows. **#2: Achieving contextual integrity.** Second, what drives the observed homogeneity in privacy policy and protocols? In particular, we would like to identify design challenges that limit the creation and implementation of bespoke privacy policies (e.g., lack of time, need, awareness, or knowledge) that reflect communities' expectations and values. Identifying these challenges will highlight opportunities for designing tools that aid customization of policies and practices toward contextually-attuned privacy protections, helping to realize one promise of decentralization. **#3: Coordinating policies across servers.** Third, we are interested in exploring design and policy issues that arise from heterogeneity in policy use. Specifically, what additional coordination between servers becomes required as data travels across them? Such coordination may involve developing methods to accurately convey information flows to users as they decide what to post to which instances, or facilitating workshops among server owners to generate cross-server privacy norms. Our work, along these three dimensions, aims to realize the potential of decentralized platforms for agency in governance and privacy through an examination of the policy work and understandings of users.

REFERENCES

[1] Cindy Cohn and Rory Mir. 2022. The Fediverse Could Be Awesome (If We Don't Screw It Up). https://www.eff.org/deeplinks/2022/11/fediverse-could-be-awesome-if-we-dont-screw-it.

[2] Evelyn Douek. 2022. Content Moderation as Systems Thinking. https://doi.org/10.2139/ssrn.4005326

[3] Ingrid Lunden. 2022. How Mastodon Is Scaling amid the Twitter Exodus. *TechCrunch* (Dec. 2022).

[4] Helen Nissenbaum. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life.* Stanford University Press. https://doi.org/10.1515/9780804772891

[5] Aravindh Raman, Sagar Joglekar, Emiliano De Cristofaro, Nishanth Sastry, and Gareth Tyson. 2019. Challenges in the Decentralised Web: The Mastodon Case. In *Proceedings of the Internet Measurement Conference (IMC '19).* Association for Computing Machinery, New York, NY, USA, 217–229. https://doi.org/10.1145/3355369.3355572

[2] As highlighted in discussions on Mastodon: https://web.archive.org/web/20230220231429/https://github.com/mastodon/mastodon/issues/23551