



# Trouble in Paradise? Understanding Mastodon Admin's Motivations, Experiences, and Challenges Running Decentralised Social Media

ZHILIN ZHANG, University of Oxford, United Kingdom

JUN ZHAO, University of Oxford, United Kingdom

GE WANG, Stanford University, United States

SAMANTHA-KAYE JOHNSTON, University of Oxford, United Kingdom

GEORGE CHALHOUB, University College London, United Kingdom

TALA ROSS, University of Oxford, United Kingdom

DIYI LIU, University of Oxford, United Kingdom

CLAUDINE TINSMAN, University of Oxford, United Kingdom

RUI ZHAO, University of Oxford, United Kingdom

MAX VAN KLEEK, University of Oxford, United Kingdom

NIGEL SHADBOLT, University of Oxford, United Kingdom

Decentralised social media platforms are increasingly being recognised as viable alternatives to their centralised counterparts. Among these, Mastodon stands out as a popular alternative, offering a citizen-powered option distinct from larger and centralised platforms like Twitter/X. However, the future path of Mastodon remains uncertain, particularly in terms of its challenges and the long-term viability of a more citizen-powered internet. In this paper, following a pre-study survey, we conducted semi-structured interviews with 16 Mastodon instance administrators, including those who host instances to support marginalised and stigmatised communities, to understand their motivations and lived experiences of running decentralised social media. Our research indicates that while decentralised social media offers significant potential in supporting the safety, identity and privacy needs of marginalised and stigmatised communities, they also face considerable challenges in content moderation, community building and governance. We emphasise the importance of considering the community's values and diversity when designing future support mechanisms.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **Empirical studies in collaborative and social computing**; **Social media**; **Computer supported cooperative work**.

Additional Key Words and Phrases: decentralisation, social media, citizen-powered internet, admin experiences, content moderation, online communities

## ACM Reference Format:

Zhilin Zhang, Jun Zhao, Ge Wang, Samantha-Kaye Johnston, George Chalhoub, Tala Ross, Diyi Liu, Claudine Tinsman, Rui Zhao, Max Van Kleek, and Nigel Shadbolt. 2024. Trouble in Paradise? Understanding Mastodon

Authors' Contact Information: [Zhilin Zhang](#), [zhilin.zhang@cs.ox.ac.uk](mailto:zhilin.zhang@cs.ox.ac.uk), University of Oxford, United Kingdom; [Jun Zhao](#), [jun.zhao@cs.ox.ac.uk](mailto:jun.zhao@cs.ox.ac.uk), University of Oxford, United Kingdom; [Ge Wang](#), [gew@stanford.edu](mailto:gew@stanford.edu), Stanford University, United States; [Samantha-Kaye Johnston](#), [samantha-kaye.johnston@cs.ox.ac.uk](mailto:samantha-kaye.johnston@cs.ox.ac.uk), University of Oxford, United Kingdom; [George Chalhoub](#), [g.chalhoub@ucl.ac.uk](mailto:g.chalhoub@ucl.ac.uk), University College London, United Kingdom; [Tala Ross](#), [tala.ross@cs.ox.ac.uk](mailto:tala.ross@cs.ox.ac.uk), University of Oxford, United Kingdom; [Diyi Liu](#), [diyi.liu@oii.ox.ac.uk](mailto:diyi.liu@oii.ox.ac.uk), University of Oxford, United Kingdom; [Claudine Tinsman](#), [claudine.tinsman@cs.ox.ac.uk](mailto:claudine.tinsman@cs.ox.ac.uk), University of Oxford, United Kingdom; [Rui Zhao](#), [rui.zhao@cs.ox.ac.uk](mailto:rui.zhao@cs.ox.ac.uk), University of Oxford, United Kingdom; [Max Van Kleek](#), [max.van.kleek@cs.ox.ac.uk](mailto:max.van.kleek@cs.ox.ac.uk), University of Oxford, United Kingdom; [Nigel Shadbolt](#), [nigel.shadbolt@cs.ox.ac.uk](mailto:nigel.shadbolt@cs.ox.ac.uk), University of Oxford, United Kingdom.



This work is licensed under a Creative Commons Attribution-NonCommercial International 4.0 License.

© 2024 Copyright held by the owner/author(s).

ACM 2573-0142/2024/11-ART520

<https://doi.org/10.1145/3687059>

Admin's Motivations, Experiences, and Challenges Running Decentralised Social Media. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 520 (November 2024), 24 pages. <https://doi.org/10.1145/3687059>

## 1 Introduction

In the evolving landscape of social media, decentralised platforms [5, 10] have emerged as viable alternatives to traditional, centralised networks. Mastodon [77], an open-source, federated social media platform, is at the forefront of this shift, championing user autonomy and community governance [94, 125]. Unlike traditional mainstream platforms governed by singular corporate entities, Mastodon's decentralised nature allows ordinary citizens to create their own social media services, enabling the creation of diverse, independent communities, each with its own culture and rules [58]. This structure not only democratises content and interaction but also introduces new dynamics in social media governance. Since October 2022, there has been a significant rise in public engagement with Mastodon, partly influenced by a migration of users from Twitter/X<sup>1</sup> [55]. Mastodon's increasing popularity underscores a growing public interest in more equitable and community-focused online spaces, challenging the dominance of centralised social media giants.

While extensive research has explored traditional social media platforms [6, 22, 83, 108], the realm of decentralised social media, notably Mastodon, remains less examined. Previous studies on decentralised platforms have primarily focused on technical aspects of the infrastructures themselves [30, 45, 119], privacy policies [58], and content moderation practices and challenges [8, 50, 51, 94], often employing quantitative methods to analyse these aspects on a large scale. However, there exists a notable gap in the literature, specifically in providing an understanding of the key stakeholders that make Mastodon work: the perspectives and experiences of Mastodon instance operators. This insight is critical for us to understand how these key network administrators experience this alternative form of social networking online, and to identify the labour and support required to sustain community building and well-being in these emerging decentralised spaces. To address this gap, in this paper we interviewed Mastodon administrators to provide an understanding of their motivations for transitioning to Mastodon, and their real-world experiences and challenges.

We first conducted a pre-study survey with 33 respondents for a broad perspective, and then followed up with semi-structured interviews with 16 participants. By doing so, the study seeks to shed light on the often-unseen efforts behind maintaining such platforms, and how these efforts contribute to the broader landscape of decentralised social media. More specifically, we asked the following research questions:

- **RQ1:** What motivated Mastodon administrators to establish and maintain their instances, and what are their perceived key benefits of decentralised social media?
- **RQ2:** What are the key challenges faced by Mastodon administrators in managing decentralised social media instances, and how do these challenges impact their experience?
- **RQ3:** What strategies do Mastodon administrators employ to navigate the challenges associated with decentralised social media management?
- **RQ4:** Based on their experiences and needs, what improvements and prospects for Mastodon and decentralised social media do Mastodon administrators identify?

Our research offers a systematic and qualitative examination of Mastodon administrators' experiences, shedding light on the unique challenges and opportunities arising from the decentralisation of social media. These insights are crucial to understanding the long-term sustainability and destiny of these decentralised platforms; will they grow and eventually replace their traditional counterparts, or will they co-exist with them (and if so, for what purposes), or will they disappear entirely? We

<sup>1</sup>The company, originally Twitter, rebranded as X in July 2023; we refer to it as Twitter/X for clarity.

find that decentralisation empowers users with greater control, yet intensifies the burden of management. Key findings reveal three primary motivations for administrators: creating safe, inclusive environments; offering personalised, algorithm-free user experiences; and enhancing users' privacy. However, these benefits are counterbalanced by increased challenges and responsibilities, such as user adoption, content moderation, and managing power dynamics. Our findings underscore future opportunities and key considerations for improving Mastodon and related technologies, which have significant implications for the development of decentralised, citizen-powered social media.

## 2 Background

### 2.1 Centralised Social Media Platforms

Online social services have evolved significantly since their inception. Initially, platforms like Usenet [52, 72], established in 1980, exemplified a decentralised approach to online social interactions. Usenet operated as a distributed network, without a central server or administrator, facilitated the free exchange of ideas while limiting content moderation enforcement [91]. Over time, the internet oversaw a shift towards centralisation, exemplified by platforms like Twitter and Facebook, which consolidated control over user interaction and data [54].

Content moderation has become one critical challenge for the centralised platforms, particularly concerning fairness and transparency [44]. The algorithms behind these platforms can inadvertently perpetuate biases [9, 103], resulting in an under-representation of diverse communities and hindering community building and diverse interactions online [110]. While centralisation allowed for the use of sophisticated algorithms for content curation and recommendation, enhancing user engagement, they also give rise to concerns about algorithmic manipulation and the formation of echo chambers [25, 100]. Algorithmic manipulation is shown to potentially exacerbate filter bubbles [38, 65, 96, 122], restricting users' exposure to diverse viewpoints and yielding polarisation.

At the same time, the rise of monetisation and advertising practices of private companies has led to the prioritisation of profit over user experience, inundating users with intrusive ads and incentivising clickbait content [63, 76]. As a result, centralised platforms can amass vast quantities of personal data, posing risks to user privacy [20, 56, 66, 75], and raising questions about data autonomy and transparency [39, 98, 115, 124]. Furthermore, the lack of interoperability [15] between platforms could limit user choice and data portability [59], hindering the ability to switch platforms or merge data from various sources.

Marginalised communities face particularly heightened vulnerability on centralised social media platforms. Prior work has highlighted content moderation policies on these platforms are often inconsistent [87–89] and disproportionately target marginalised communities [49]. Studies have found that user groups such as political conservatives, transgender individuals, and black people experience more frequent content and account removals [48]. This, coupled with the prevalence of harassment and hate speech on these platforms, creates unsafe environments, silencing marginalised voices [79, 82, 111, 126]. Finally, privacy and safety concerns, such as data breaches and misuse of personal information [14, 40, 47], can often have dire consequences for LGBTQ+ individuals, in particular who may not yet be out about their identity [46, 85, 92].

### 2.2 The Potential of Decentralised Social Media

Decentralised social media represent a paradigm shift in the realm of online communication [5, 10]. Unlike their centralised counterparts [68] that are mostly governed by a single private entity, decentralised platforms operate on distributed or federated networks and protocols [30]. In a decentralised framework [45], the control over data, content moderation, and network infrastructure is dispersed among a multitude of nodes or participants, reducing the dominance of any single

authority [69, 70, 112]. This alternative structure can foster a sense of user ownership and autonomy, offering an alternative approach to online interactions [121, 125].

Contrary to a centralised platform, a decentralised social media network is connected by a common protocol, thus users are not tied to a single hosting server, allowing them to be connected across the network. One notable example is the Fediverse [70], underpinned by the ActivityPub [113] protocol. Other platforms, e.g. diaspora\* [2] and HubZilla [3], also explore alternative protocols for various purposes, with or without bridges to the fediverse. Prominent decentralised platforms in the fediverse encompass microblogging services like Mastodon [77] and Pleroma [93], along with video-sharing platforms such as PeerTube [90]. Notably, Elon Musk's acquisition of Twitter [27] further accelerated the surge in popularity of the Fediverse, leading to a substantial migration of Twitter/X users to these decentralised platforms [60, 118].

Numerous studies have examined the design and structure of decentralised social media platforms, especially Mastodon. Zignani et al. [119, 120] analysed Mastodon's physical architecture and instance-level evolution and how the design has been facilitating user connection. La Cava et al. [70] similarly explored connections between Mastodon instances. Beyond social networking, Doan et al. [33] assessed the performance of a decentralised video streaming application. Raman et al. [94] examined re-centralisation challenges for decentralised social media. These studies characterise key technical elements of decentralised platforms. Moreover, recent research sheds light on these platforms' potential to protect privacy [109]. In essence, decentralised social media has the potential to offer solutions to various problems pertinent to centralised platforms and address concerns at large [74]. In the following section, we will introduce Mastodon as an example to explore the possibilities of decentralised social media.

### 2.3 Introduction to Mastodon

Mastodon, introduced in 2016, represents a pivotal development in the landscape of decentralised social media platforms [17, 94]. As one of the largest and most popular Fediverse projects [1], Mastodon challenges the conventional centralised social media model by embracing user autonomy, data ownership, and community-driven content moderation principles [125]. Mastodon operates on open-source software and relies on a federated network comprising independently administered servers known as "instances" [78], interconnected through a standard protocol, ActivityPub [113]. These interconnected instances facilitate seamless interactions among users while retaining local moderation authority. As the decentralised social media movement gains momentum [55], Mastodon emerges as a prominent case study worthy of examination within the broader discourse on the future of social media.

**2.3.1 Content Moderation and Governance.** Content moderation represents a vital aspect of online social platforms. Facebook and Twitter employ hierarchical, commercial content moderation systems with paid moderators, adhering to a universal platform policy [43]. Unlike most centralised platforms with top-down content governance mechanisms [97], Mastodon adopts a decentralised model [8, 50, 51] where local instance administrators (and selected moderators) play a key role in content moderation and determine moderation rules (e.g., community guidelines) for their respective instances [125].

Content moderation has been extensively researched for centralised platforms, such as Reddit. Both Mastodon and Reddit are community-driven social platforms [34], and focus on fostering smaller, more niche communities. Hwang and Foote [57] have shown that small online communities play unique roles in providing specific informational and interactional spaces. Reddit utilises a mix of official platform policies and volunteer moderators within its sub-communities (subreddits) for localised governance [23, 37]. In contrast, Mastodon operates with greater decentralisation,

where independent servers (instances) follow a shared covenant [4] without a central authority, resembling Reddit's structure of autonomous sub-communities but with more independence [84].

Comparing Mastodon and Reddit, Nicholson et al. [84] showed that Mastodon's community rules prioritise addressing harassment and hate issues more than those on Reddit, potentially due to issues encountered on other platforms and reflecting insufficient support for instance administrators. Extensive prior research has also explored moderation challenges on Reddit. For example, Gilbert [42] identified that visible moderation is often perceived as censorship and reflects default masculine whiteness, which complicates fostering empathy and inclusivity within the community. However, while prior research has proposed design frameworks for decentralising platform power in content governance [24, 61], few empirical evidence examines downstream impacts across different user groups on decentralised platforms.

**2.3.2 Feed Control and Agency.** Social media feed experiences have become concerns within centralised platforms, where proprietary algorithms often prioritise content for user engagement and profit, potentially perpetuating issues such as biases and echo chambers [13, 25]. In contrast, Mastodon feeds differ from other social media platforms primarily in their non-engagement-based, strictly reverse chronological content presentation [36]. This approach allows users greater control over their feed content, aligning content directly with their chosen networks and interests rather than algorithmically suggested posts. Mastodon uses a non-profit financial model, relying on grants and crowd-sourced funds rather than commercial monetisation strategies [12, 78]. While this approach allows better alignment with user interests, it could limit resources available for creating refined, widely appealing experiences [21, 36].

**2.3.3 Data Privacy and Autonomy.** Data privacy and autonomy constitute another pivotal consideration for social media platforms dealing with user data. Centralised social media platforms collect and track vast amounts of user behaviour data to sustain their advertising and content targeting business models [31, 106, 123]; as a result, they offer limited privacy options and could expose users to privacy violations [95], data breaches, and potential forms of data misuse [73, 80, 104]. In contrast, Mastodon disperses data across numerous independent servers rather than centralising it in a single entity [45, 78], reducing the risks of mass data collection and misuse, as no single server holds extensive information about all users. Being open-source and community-driven, Mastodon enables server administrators to tailor privacy policies aligning with individual or collective community principles [58].

**2.3.4 Challenges for Decentralised Social Media.** Moderation and governance pose a major challenge for decentralised social media. The growing popularity of the Fediverse challenges administrators with increasing moderation burdens, especially in larger instances [51], and users are reliant on the administrative team behind each instance [5]. While instance-level control allows tailored content moderation policies aligned with community needs, the inconsistencies and inefficiency may undermine inclusivity and safety as the user base scales [50]. Instance-based moderation in decentralised platforms could result in "collateral damage," where non-harmful users are blocked due to the actions of a few, potentially leading to user dissatisfaction and attrition from the platform [51]. Moreover, while decentralised platforms provide opportunities for customised privacy policies, recent research showed a prevalent use of the same generic policy texts, prompting concerns about the alignment with user preferences, and the challenges of coordinating policies across diverse communities [58]. Prior research [94] indicated that while Mastodon sees active engagement across diverse instance types, it faces pressures of re-centralisation. Additionally, there is a growing need for research into distributed AI methods to enhance moderation and address potential threats from other nodes [81]. Recent research showed that decentralised social media platforms struggle with

managing illegal content, specifically child sexual abuse material (CSAM), due to the challenges of adapting traditional safety tools designed for centralised networks to their unique, decentralised architectures [107].

### 3 Methods

The study began with a pre-study survey, followed by semi-structured interviews with 16 Mastodon administrators across 7 countries (USA, Australia, Costa Rica, Germany, The Netherlands, France, and Canada), to obtain exploratory insights into the overarching experiences of Mastodon administrators. We aimed to understand participants' motivations, challenges, strategies, and future prospects for decentralised social media. The study took place between August 2023 and January 2024, and was approved by the Institutional Ethics Review Board at our university.

#### 3.1 Participants Recruitment

To recruit participants, we visited the Join Mastodon website<sup>2</sup> to send invitations and a preliminary survey to instance administrators. We reached out to instance administrators covering a range of general topics, as well as those focusing on supporting marginalised and stigmatised communities. We also posted general recruitment advertisements on social media platforms. Alongside this, we employed snowball sampling, recruiting a few participants from the recommendation of participants who had taken part in our study.

In the preliminary survey, respondents were asked if they were willing to proceed with our interview phase. Those who agreed were contacted via email to schedule an interview session. We successfully scheduled and conducted interviews with 16 participants. In preparation for the interview, participants were sent an information sheet and a consent form, which they needed to complete and return before the interview could take place. Once we received their consent, we conducted remote, semi-structured interviews over Zoom. These sessions were audio-recorded for analytical purposes, with the consent of the participants.

#### 3.2 Preliminary Survey

We had a total of 33 respondents completed the 15-minute preliminary survey. This survey was designed for the recruitment of semi-structured interview participants and to facilitate preparation for the interviews. This survey enabled us to determine whether prospective participants met the inclusion/exclusion criteria. To participate in our study, participants must be at least 18 years old and administer at least one Mastodon instance. Further questions from the survey permitted us to obtain contextual details about prospective participants, including the length of time they have been an administrator, the number of local users within their biggest instance and the intended topics of their instances. Additionally, we asked open-ended questions, including their motivations for starting the instance, the values underpinning the instance, their challenges of administering the instance, and views on the future of decentralised social media. The survey provided a sufficient sample for gathering initial perceptions and informed the design of our follow-up interviews.

#### 3.3 Semi-structured Interviews

We conducted semi-structured interviews with 16 Mastodon administrators, who consented to participating in the interviews and met all our recruitment criteria. Table 1 details the demographic information of the interview participants. We received a diverse sample composition with instance topics including General, LGBTQ+, Disability, and Activism. In the interviews, we delved into participants' experiences as an administrator for their Mastodon instance(s). We started by asking

---

<sup>2</sup><https://joinmastodon.org/>



Participant ID	Topic of instance	Number of instance users	Tenure of instance administration	Continents of instance administrator
P1	LGBTQ+	100 - 999	6 - 12 months	Oceania
P2	LGBTQ+	1,000 - 9,999	More than 12 months	North America
P3	LGBTQ+	1,000 - 9,999	6 - 12 months	North America
P4	LGBTQ+	100 - 999	6 - 12 months	North America
P5	Disability	10 - 99	More than 12 months	Europe
P6	LGBTQ+	100 - 999	6 - 12 months	Europe
P7	General	Over 100,000	More than 12 months	Europe
P8	General	10,000 - 99,999	More than 12 months	Europe
P9	LGBTQ+	0 - 9	6 - 12 months	North America
P10	General	10 - 99	More than 12 months	North America
P11	LGBTQ+	10 - 99	6 - 12 months	North America
P12	General	1,000 - 9,999	More than 12 months	Europe
P13	Activism	1,000 - 9,999	More than 12 months	North America
P14	General	100 - 999	6 - 12 months	North America
P15	General	10,000 - 99,999	More than 12 months	North America
P16	General	1,000 - 9,999	More than 12 months	North America

Table 1. A summary of participants along with basic information about the communities they support.

their motivations for setting up their Mastodon instance. Subsequently, participants were prompted to reflect on the challenges associated with running the instance. After that, they were prompted to elaborate on the strategies they employed to address or mitigate such challenges. Finally, participants were invited to share their thoughts on potential improvements and future prospects for Mastodon and decentralised social media. On average, each interview session lasted for 1 hour. Each participant was compensated with a £12 Amazon.com gift card to thank them for their time and participation in both the survey and interview components of the study.

### 3.4 Data Collection and Analysis

We transcribed the audio recordings of the interviews and stored the transcripts in a password-protected folder on the university system. All personally identifiable information has been redacted in the transcription to protect the privacy of our participants. We conducted a grounded, thematic analysis on the transcripts [11, 16, 18, 19]. The thematic coding process started by dividing the transcriptions into two equal-sized sets. Two authors independently analysed the first set of the transcription to derive an initial set of codes. They then met to consolidate codes into a common codebook, with a Cohen kappa of 0.92. The first author then used this codebook to code the remaining set of transcription. As this is an exploratory study, the sample we reached out has provided a good range of themes relevant to the scope of our study. Given the exploratory nature, our primary focus is to gain an in-depth qualitative understanding of the experiences, perceptions and barriers faced by Mastodon administrators. Therefore, rather than pursuing a large-scale study, we believe the sample size is sufficient to meet the objectives of this research.

### 3.5 Research Ethics

This study was approved by the Institutional Ethics Review Board at our university. In Table 1 we listed continents instead of countries to safeguard the anonymity of participants while still providing valuable geographical context. By opting for continent-level description, we seek to maintain the integrity of the data while reducing the potential for inadvertent re-identification of individuals within specialised communities.

## 4 Results

### 4.1 Motivations for using Mastodon and Launching the Instances

**4.1.1 Creating a Safe and Inclusive Online Space.** The main motivation for most participants to transition to Mastodon, from traditional social media, is to feel safer online and better supported for their identity and by their community. A lot of these feelings stemmed from marginalised and stigmatised communities, such as LGBTQ+, disability, and activism communities, who expressed

having faced harassment and feeling a lack of safety and representation on traditional platforms. For example, P13, admin of an instance with the topic of activism in lifestyle, shared their past experience of encountering harassment on traditional social media platforms. They mentioned receiving disturbing images when they posted anything publicly with an activism message, while little moderation actions would be taken; in contrast, on their Mastodon instance, they had more opportunities to “ban that user because they’re not welcome on our server,” which made them feel that they “would have some place that we feel safe”.

The majority of the participants expressed that their need for a safe space is largely attributed to traditional platforms’ insufficient support for diverse user values and identities. For example, P2, an LGBTQ+ instance admin, shared the following:

I already do not feel safe or comfortable with Twitter<sup>3</sup> as a platform [...] You know, all the social media platforms are global, or at least nearly global, but they try to serve the lowest common denominator, so they are unlikely to make moderation decisions that could alienate a large portion of their user base. (P2, LGBTQ+)

Participants’ sentiment was echoed in the language they used; many referred to their Mastodon instance as a “home,” “bubble,” “island,” or “haven,” underscoring their deep sense of safety and belonging within their instance, a priority repeatedly emphasised by many participants, as articulated below by P4, an LGBTQ+ instance admin:

[I moved to Mastodon] because when I was on Twitter, I noticed that there was a hostile environment towards minority communities. So when I started my Mastodon server, I wanted to be sure that this is a safe space for LGBTQ+ people and minority communities. It’s an open place for everyone while you respect (others) and don’t engage with hate speech and things like that. It’s intended to be a safe space. (P4, LGBTQ+)

Participants highlighted that fostering safer, more inclusive spaces on Mastodon enables marginalised users to discuss aspects of their identity that they might otherwise feel unsafe or uncomfortable doing so on traditional platforms:

It (Mastodon) has been a welcoming space where I can feel more free to be myself. By that I mean, I can talk more about my personality, I can talk more about my struggles that I have on other social media, I can talk about my marginalised status, you know, as somebody who is queer, somebody who is trans, somebody who is polyamorous. (P2, LGBTQ+)

A large number of participants also highlighted that safety was also associated with Mastodon’s features of allowing admins to customise moderation according to their community needs. For example, as P2 shared, “Facebook might say, we’re not going to ban somebody because it’s going to reduce the number of users and we’re focused on growth. In contrast, Mastodon allows admins to employ a different moderation strategy that is more customised to their community needs [...] We make a decision on behalf of our members, and we try to focus on who will be the most harmed, as opposed to what decision will give us the most people.”

**4.1.2 Empowerment Through Mitigating Algorithmic Manipulation.** Another main motivation for many of the participants to transition to Mastodon is to avoid algorithmic manipulation associated with traditional social media platforms. For example, P13 shared that,

I didn’t like the algorithms on Reddit, Facebook, or Twitter. It seemed like you would have to fight a lot against them in order to build an audience and to generate discussion or get anything done [...] I really wanted to get away from the walled garden of Reddit and the algorithms that control what gets attention and what doesn’t [...] I find that social networks are meant to be addictive. They try to keep people’s attention [...] I just see them as like addiction machines to produce financial gains for the companies. (P13, Activism)

<sup>3</sup>Despite the company’s recent rebranding to X, many participants used the familiar name, Twitter, during the interviews.



To the contrary, P13 preferred Mastodon as “a system where server administrators are disincen-tivised from launching advertisements or predatory algorithms.” P1 (an LGBTQ+ instance admin) described that “one of the benefits (of Mastodon) is there’s no algorithm to keep you on for as much time as possible. It’s what you follow is what you see.”

Many participants compared Twitter/X and Mastodon’s timelines and pointed out that they preferred Mastodon’s timeline for its algorithm-free experience and better usability:

(On Twitter/X) the amount of advertisements were increasing [...] they just started injecting more and more things that wasn’t something I was following [...] So, Mastodon is kind of a nice reset to how Twitter used to work, where people would post things and you see their posts and there wasn’t really any advertising or a bunch of suggestions and algorithms and things being shown that you don’t really care about necessarily. (P11, LGBTQ+)

More specifically, many participants mentioned that they preferred the ability to control their social media timeline, instead of being pushed by complex recommendation algorithms:

I really don’t want there to be some algorithm that I don’t understand. That is controlling what lands in my feed. I would much rather be in control of that and to be able to say things like no boosts from this person. And that level of control of my timeline is very pleasant. (P16, General)

Some participants shared that an algorithm-free system such as Mastodon enables them to have better social interactions, while platforms with complex recommendation algorithms could inhibit meaningful engagement:

They (algorithms) seem to inhibit engagement more than enable it in my experience [...] the algorithms are limiting, and favouring other things [...] Mastodon system doesn’t use algorithms. It uses human interest and gestures to federate content [...] It is more effective than algorithms. (P14, General)

In addition, a few participants mentioned how the algorithms could affect their digital mental health, for example, P1 shared “since I have ADHD, it’s pretty easy to be addicted to things [...] it (algorithm) is just not good for my mental health.”

**4.1.3 Advocacy for Data Privacy and Autonomy.** One last issue commonly mentioned by participants is the concern with their data privacy on traditional social media platforms, for example, how their data may be misused, shared with third-party platforms, or how some of their sensitive posts may be promoted to the public or friends unintentionally, as mentioned by P2 in the following:

A centralised social media platform does whatever it can to gather as much data about a user as possible, including purchasing them from third-party sources [...] It’s much easier for a company like Twitter to associate your accounts together. Twitter knows that you have accounts A and B. The problem is part of the algorithm is on larger social media networks will do things like recommend your private profile to your public, friends and family because they know that it’s you for both of those accounts. And you might not want that. (P2, LGBTQ+)

This data sharing and insufficiency of privacy can be particularly worrying for users from marginalised communities. For example, P5 (admin of an instance focused on disability) mentioned how such data sharing prompted their Facebook and WhatsApp abandonment due to privacy concerns: “I stopped using Facebook since they acquired WhatsApp. The sharing of information between the services is a problem for me to know that they are using my phone number to target me in other services [...] this is not acceptable for me.”

Many participants find decentralised social media can better support their data privacy and autonomy, for various reasons, including the ability to 1) decide how their data is managed, e.g. “everyone can start up their own server, so they have total control over where their data resides” (P7, General); 2) choose where their data can be stored, as mentioned by P14, who emphasised the importance of selecting a GDPR-compliant jurisdiction; and 3) freely transfer data and social

connections from one platform to another, for example, as described by P4, “You cannot take your connections with you (on traditional social media) [but] the Fediverse and Mastodon is better for autonomy. Because I can bring all my connections, like my followers, and people they follow to a different place.”

## 4.2 Challenges related to Running a Mastodon Instance

**4.2.1 User Adoption and Community Building Challenges.** Many admins shared that, decentralisation as a new concept could be difficult for users to understand, which hinders adoption. As P7 shared, “They don’t understand the concept. Sometimes they don’t understand what is a server? Why are there different servers? Which one should I pick? For some users, that’s just too much.”

This learning barrier made it difficult to onboard new users, especially those without sufficient technical background, as P5 shared, “There are a lot of people argue that’s too complicated [...] It’s a problem with a nerdy software like Mastodon [...] when it comes to (novice) users that are new to a network like that, it can be intimidating.”

Simultaneously, another major challenge mentioned by many participants was that, since many Mastodon instances are often new and small in the number of users, the amount of content on these instances could be sparse. As a result, a new user on an instance would often find their home feeds too “empty” and fail to provide a sufficiently engaging experience for many users. As P4 shared, “The users feel like it is empty. They feel quiet, like, nobody is here”.

In addition, there are barely any recommendation algorithms on Mastodon to suggest content and engage users. While this may appear as a positive consequence for some, others who are already accustomed to traditional social media platforms could feel Mastodon is too bland and lacks engaging content:

Platforms like Twitter and Facebook are really good at recommending nice content because they want users to be engaged [...] Sometimes users can feel that Mastodon is not as funny as, or as engaging as, Twitter, Facebook or whatever. I think that that’s a big challenge. (P4, LGBTQ+)

A few participants mentioned that trust can be a further barrier to user adoption of a citizen-run platform compared to platforms run by established organisations or companies. For example, as P4 shared, “Twitter or Facebook have to comply with laws [...] you trust those companies because the governments are watching them. But how can someone trust me because I am a small operator over the internet? And I think that’s a challenge [...] users have to trust someone, and they don’t know who is trustworthy and who is not.”

**4.2.2 Moderation and Governance Challenges: Balancing Individual and Community Needs.** A dominant challenge mentioned by all participants was the complexities of content moderation and community governance. While Mastodon empowers admins to customise moderation for their community, we observed a range of complex challenges that an admin must navigate in order to ensure the identity and plurality of the community are well respected.

To begin with, while admins now have more autonomy to control the content posted on an instance, most of them often mentioned the increased challenges related to balancing between freedom of speech and creating a safe, inclusive environment within the instance:

Sometimes when we do moderation, a user might say what they posted wasn’t so bad, or we’re censoring their political views in some way. For example, there are different political views on things like COVID vaccines. People might post an essay that is sceptical of COVID vaccination. And we have to moderate against such things because this can be some sort of public health issue to spread misinformation. But these are political views. Censoring these things can be a bit beside our position, or it might be a bit arbitrary. (P13, Activism)

Furthermore, “creating a moderation policy that everybody likes” can be challenging, as P13 shared. For example, when they received a suggestion to moderate against fat phobia in the community, they felt this was challenging, as many members of the community often discuss weight loss and diet benefits, which could inadvertently alienate people who experience fat phobia. Implementing rules against such behaviours requires careful consideration, as it involves balancing the need for a safe, inclusive space with the risk of penalising otherwise rule-abiding users.

In addition to moderating the content within their instances, many admins mentioned the challenge they had to face when having to moderate content from other instances. While it may be straightforward to forbid content from disruptive instances, such a decision is often more nuanced and requires a careful balance between openness and safety. For example, as P11 (an LGBTQ+ instance admin) shared, “Even though there are servers that have bad people on them, there are also LGBTQ+ members on those servers. So if we just blocked those servers, then we’re cutting off members of the community.”

The issue is further exacerbated when instances have different content moderation policies, varied values and community development maturity. The latter could particularly affect the clarity of the content permitted on an instance, and may lead to misaligned expectations of what content is appropriate on one instance versus another. Many admins found this lack of coherence could be difficult to manage rule incompatibility across Mastodon instances, as P6 shared,

Some instances have stricter rules, and some have looser rules. If the rules are too loose, we have to consider if this instance is too incompatible with us? [...] And if they’re too strict, we have to find the middle ground [...] So the difficulty is finding the middle ground between those rules, figuring out what’s acceptable, maybe even talking to the admins of that instance. (P6, LGBTQ+)

Finally, some participants have concerns that when Mastodon admins have more control over their instances, this control can come with more power and responsibility. Participants are worried that some admins have the risk of abusing the power and making moderation decisions that can negatively impact both members of their instances and people in other instances:

There have been a few high-profile instances that had rough members of their (moderation) team [...] (their admins) being super too proactive with server bans, and they decided to block the entire server [...] Those experiences have been extremely tough for both users on those instances, as well as people on our instances who are friends with people on those instances [...] It is a risk to have so much power in hand. (P11, LGBTQ+)

**4.2.3 Burdens for Administrators.** Many admins mentioned how they are now facing increased burdens of addressing scams and verifying claims on Mastodon, as mentioned by P7: “There’s a lot of people trying to scam you on Mastodon [...] That’s a very difficult situation because you don’t want your users being scammed, but you also don’t want to block users who really are in difficulties and need money.”

Several participants also mentioned the stress and psychological impact of moderation. Currently a majority of the moderation is done manually, which could be tedious and stressful for the moderators. In addition, there is a risk of being exposed to harmful content. For example, P4 shared how the process of moderating harmful content could be traumatic for the moderators: “I was exposed to the content of child abuse. And it was a hard day.”

A few participants mentioned things can get personal, and moderators can experience harassment because people know who is moderating. As P2 shared, “Since people know who was doing the moderation, they can form a lot of opinions based on your personal postings. And they can also individually harass you by that.”

Furthermore, a small number of participants mentioned legal and compliance issues that are unclear for running a citizen-powered platform, as P4 shared,

How do you identify very bad servers legally? Like servers that host child pornography, for example, you cannot look for that. I want to find those servers so I can block them, but that could be an illegal activity. It depends on your jurisdiction. [...] It's a challenge because I'm not an expert in law. So I'm not really sure what are my legal implications in my country. (P4, LGBTQ+)

**4.2.4 Navigating New Ways of Privacy Protections.** A few participants mentioned minor concerns regarding data privacy, mostly surrounding uncertainty of where the data ends up, and how to get rid of it:

I know that there's no one central repository of data about a user within the fediverse [...] Content is disseminated via the federated connections that your server has and then cached or downloaded and stored on many different servers [...] You don't know where your information ends up, and you can't get rid of your data because it's everywhere. (P10, General)

Regarding privacy policies, most participants indicated that they either retained the default policy, made only minor modifications, or adopted policies created by others. While this approach may deviate from one of the core principles of decentralised social media, which provides the flexibility to customise privacy policies for each instance, many participants seemed uncertain how much their users may be interested in reading the details of privacy policies:

I don't think I modified anything with mine. I just Googled a few other instances that have privacy policies, and that's basically where I got it from [...] I can probably guarantee you, not a single one of my users have actually gone there to check the privacy policy. It's not something that people just genuinely do. Unfortunately, you know, they often don't read the privacy policy. (P3, LGBTQ+)

### 4.3 Strategies for Running a Mastodon Instance

**4.3.1 Prioritise Quality over Growth: Keeping Instances Small.** A prominent theme, especially among participants who manage instances for marginalised or stigmatised communities, centred on their approach of prioritising the quality of the instance over its growth. Many participants mentioned that they would rather keep their instance small<sup>4</sup> to cope with the content moderation and governance challenge:

I think content moderation is just a hard problem in general. I don't think there's a quick fix for it. I think the best way to do it is to have a small enough community that you kind of know everybody or at least kind of recognise them by name. And you can trust them to not be like bad actors or doing something in bad faith. (P9, LGBTQ+)

And to ensure a safe and friendly community:

It (my instance) was a smaller, more cosy space [...] It doesn't have as broad of a reach so the people who choose to be there understand the community [...] (Being small) would eliminate trolls [...] Smaller communities, decentralised communities, overall do a much better job of moderation and protecting their users [...] We can be reactive to our small community, as opposed to a very large community. (P2, LGBTQ+)

Some participants believe that small instances are better aligned with the nature of decentralised networks, which permits more flexibility to cater for the needs of representing diverse identities, communities, and values:

I want to keep it simple and small. Because when the instance grows, your responsibilities are growing [...] It's the nature of decentralised networks. The idea is not to have one or two big instances. The idea is to have thousands, maybe someday millions of instances, each with a small number of users. Because it would show how diverse the world is with many people in the many topics out there. (P5, Disability)

<sup>4</sup>the definition of being 'small' can differ between different participants, ranging from several to thousands of users.

A few believe that a smaller space can lead to a more rewarding social networking experience, which can provide a suitable alternative for many who struggle to secure an affirmative experience on large platforms:

If you go to a party, and there are 1,000 people, you'll never have an idea about who to talk to. But if you go to a party, and there are 10 strangers you've got, there are more chances that you're going to interact because there you have less choice of people to interact with. And you might actually get a lot more out of it than you would at the larger party. And I think that could be an analogy to Twitter versus Mastodon. It's a smaller space. (P3, LGBTQ+)

**4.3.2 Collaborative Content Moderation.** Many participants indicated the presence of a dedicated moderation team within their instances. These team members are usually selected from the active user base, who are often more acquainted with the community dynamics and recognised by fellow members. These factors enhance their suitability for contributing positively to the community.

Some others described intentionally creating a diverse team of moderators with varying language abilities, cultural backgrounds, and geo-location distributions to achieve more effective content moderation:

Our moderation team right now has about 10 people. The moderation team covers a variety of perspectives and also covers a variety of time zones. We try to have good coverage where we can. Using the Japanese moderator as an example, we explicitly brought them on board because we felt the need for Japanese-specific moderation for cultural context. (P2, LGBTQ+)

In addition, some participants expressed that appointing moderators requires a significant level of trust. As P9 shared, "It would be people that I've either met in real life or have talked with, known long enough that I can sort of totally trust them [...] like someone I would give my passwords to, someone I totally trust."

However, when asked about how they communicate and collaborate with their moderation team, many sought to use separate software. For example, some preferred using "discord separate of the mastodon instance" (P2) for moderation team communication due to the absence of sufficient communication functionality on Mastodon.

In addition, most participants find it helpful to engage citizens of the instance in moderation by receiving reports from them: "I would prefer that more people report issues as they see them" (P2). Many participants further mentioned collaborations and communications with admins from other instances for "continuously communicating or helping each other" (P4). Most participants prefer a collaborative moderation style of outsourcing the moderation responsibility to mitigate their burdens. They receive help from their moderation team, instance members, and friends and allies from outside their community.

On the other hand, another motivation for doing collaborative moderation is to make sure that the moderation decisions are more democratic and to avoid personal bias and "dictatorship":

Not having other people to help me make tough moderation choices, I guess, is the hardest part and leads me towards wanting to grow the governance so that when decisions are made, they're not made by a single person [...] It is like a mini dictatorship, a benevolent dictator, and it's me. And so there are two (moderators) because I didn't like the idea of a dictator. (P14, General)

**4.3.3 Proactive and Reactive Moderation.** Some participants mentioned that they take a proactive moderation style to prevent harmful content before it could reach their instance. For example, P2 shared, "We choose to be proactive, as opposed to waiting for harm. We will work to block those people when we find them, as opposed to only waiting for people to report them." Furthermore, these proactive admins often exhibit a heightened sense of autonomy, driven by their commitment to ensuring the safety of their instance:

I am the governor of my own Mastodon instance. And like I said, it's not a democracy. I think the biggest problem is a lot of people want to treat some of these private things like democracies, but they don't have to live here and no one's forcing them to live by my server. So you know, it's my rules. (P3, LGBTQ+)

While being proactive in moderation has the benefit of enhancing the safety of the instance, it could potentially impede openness and receive criticism. P2 shared their experience, "We've had people explicitly get angry for both themselves and on behalf of other people, saying that we're being exclusionary, we're being overbearing, we are moderating above and beyond what we should."

Some admins prefer to take a reactive approach to moderation and rely more on user reports. For example, as P8 (a general instance admin) shared, "We do not make any prior moderation. We do not do active moderation. We moderate when we receive signals from users. So the user base helps us in that way. They are the ones who see issues or problems with some messages or interactions and signal them to us."

#### 4.4 Improvements and Future Prospects

**4.4.1 Supporting Users with their Onboarding Process.** A major theme of improvements mentioned by participants is user onboarding. Participants shared several possible options they would like to see for enhancing the user onboarding experiences, including "getting-started wizards", "starter kits", and "tutorial system" to help a new user get familiar with the platform. For example, P1 shared, "Having a getting-started wizard when you make an account on a Mastodon server that kind of holds your hand through it can be pretty useful."

**4.4.2 Supporting Content Discovery with Customised Algorithms.** For the challenge regarding Mastodon being too "empty," a few participants are devising their own solutions, such as manually-curated social media feeds (P15) or creating specific algorithms for the community (P13). P13 utilised the open-source feature of Mastodon and developed an activism-focused algorithm, which analyses names and bios for activism-specific indicators and then boosts content from these users, creating an activism-centric feed. This approach can potentially help new users find engaging content, fostering a welcoming and relatable environment for community members without relying on complex sentiment analysis.

However, many participants expressed reservations towards opaque recommendation algorithms. A few showed openness to transparent algorithms that are understandable and controllable by users. As shared by P10, "I am open to the idea of there being some simple recommendation algorithms added to the onboarding experience. I am wary as the more complex and opaque that gets, the more pitfalls it gives rise to."

**4.4.3 Improvements for Effective Moderation Tools.** Most participants mentioned the need for more effective moderation tools. They have used tools like Discord bots for report handling and account management (P1). In addition to that, many participants would like to have tools to help them automate some of the manual moderation process, such as automatically detecting harmful content or addressing user reports. As P4 shared, "I think it would be better to have more moderation tools like detecting spam or hate speech [...] We could have a local engine that detects new content similar to the ones reported in the past."

In addition, a few shared their desire for a meta platform to assist them with instance management, especially when the instance grows larger. As shared by P14, "I think a meta platform to help you organise and govern your instance would be helpful to be able to handle larger moderation [...] A meta platform to help people go from hobbyist to institution or organisation."



**4.4.4 Prospects for the Future of Mastodon and Fediverse.** Most participants believe that Mastodon and similar federated social media will coexist with, rather than replace, traditional centralised commercial platforms. They view Mastodon as a more privacy-conscious and ad-free alternative, especially appealing to those disenchanted with the mainstream social media's management and user experience. Participants also note that Mastodon may not appeal to those aiming for viral success, such as influencers or brands, due to its unique structure and community-focused culture. Instead, it is seen as a haven for specific interest groups, including activists or individuals seeking safe online spaces. There's a sentiment that Mastodon may continue to be a specialised tool, similar to "Linux in the operating system world." Some speculate about the possibility of commercial platforms adopting federated models or integrating with services like Mastodon to improve interoperability. Yet, participants highlighted that Mastodon's user onboarding challenges, owing to its less straightforward interface and the complexities of its infrastructure and moderation demands, may limit its growth compared to more mainstream platforms.

## 5 Discussion

### 5.1 Key Findings and Contributions

This study provides a systematic and qualitative examination of Mastodon administrators' experiences, offering insightful perspectives on the management of decentralised social media. We uncovered three key motivations driving Mastodon administrators to establish and maintain their own platforms, including creating safe, inclusive environments; offering personalised, algorithm-free user experiences; and enhancing users' privacy. However, these benefits are counterbalanced by increased responsibilities like ensuring user adoption, striking a balance among diverse preferences in content moderation, and adapting to new ways of managing data privacy.

While our findings resonate with previous research [41, 94], we observed a strong desire among our participant administrators to foster cooperative and trusting relationships within their user communities while addressing these challenges. Administrators have explored strategies such as focusing on the quality of their instances rather than sheer growth, adopting a collaborative approach to moderation to distribute the workload, and implementing proactive moderation to preempt potential harms. At the same time, administrators are in urgent need of better support for their community building and maintenance, while ensuring a cooperative and trusting community relationship. These findings provide critical contributions regarding our understanding of the challenges faced by Mastodon administrators, including those supporting many stigmatised and marginalised communities; instead of relying on supports that focus on providing effective technical performances, our findings indicate that we must consider the plurality of community values and relationships when designing solutions for maintaining sustainable, citizen-powered, decentralised social media platforms.

### 5.2 Safe Spaces and Content Moderation

The central motivation of participants to foster a safe and inclusive environment resonates with evolving concepts of safe spaces [26, 64], and particularly with previous research exploring the importance of online safe spaces for marginalised individuals and communities [7, 47, 71, 101, 105, 117]. The concept of "safe spaces" emerged as a salient groundwork organising tactic in the late twentieth-century United States through the rise in feminist, queer, and anti-racist movements [26]. Safe spaces, often through separatism, provide a level of respite from ongoing marginalisation and physical, verbal, and emotional harms, while also creating space for community building and collective activism [64]. Emerging research on online safe spaces on traditional social media platforms, in-line with the findings of this study, highlight that cultivating and maintaining safe

spaces often involves a combination of closed, small-size groups [35]; network separatism; avoiding censorship; and community-centred and contextual moderation [7, 47, 86]. Unfortunately, safety is often an ideal, yet not a concrete reality in online (and offline) safe spaces, particularly for multiply marginalised people [64, 86, 101, 114]. Previous literature highlights how, even within online safe spaces, social media can replicate and exacerbate intracommunity tensions and power dynamics [114] and harm people at the intersection [28] of multiple marginalisations [86]. These forms of intracommunity harm complicate content moderation approaches and highlight the necessity of intersectionality in approaches to safety and content moderation on Mastodon instances.

Our findings highlight how some Mastodon instances are approaching such challenges through intentionally forming diverse teams of moderators. However, greater attention is needed in future work to the ways in which both inter- and intra-community harms are addressed and multiply marginalised identities are supported on Mastodon, in-line with recent calls to start attending to intersections of identity in HCI [102]. Beyond safety and inclusion, participants expressed values of autonomy, privacy and self-determination, which aligns with previous literature on marginalised users' values and their misalignment with perceived values of mainstream social media platforms [32]. This study highlights that Mastodon presents the potential to support such user values, yet further consideration and support is needed to address anxieties over power dynamics in instance moderation and governance, highlighted both through this study and previous literature on platform governance [32].

Contrasting with previous research positing homogeneity in privacy policy adoption between Mastodon instances [58], our investigations reveal a disparate application of moderation policies. Despite superficial congruities in terms of service and privacy policies, our data suggest pronounced divergence in moderation strategies. The query arises: Why is there such variability in moderation strategies? The answer lies in their inherent diversity and the distinct expectations they embody. This diversity is emblematic of the virtues of decentralisation, which fosters a spectrum of moderation practices, including reactive and proactive methods, to engender environments perceived as safe by their communities. While decentralised networks aspire to distribute governing authority more equitably, fostering grassroots community development [61], the ongoing discourse on digital platforms highlights a significant issue – inconsistency and bias in moderation as shown in the Reddit example [116]. However, inconsistency in moderation presents less of a conundrum for smaller instances on Mastodon compared to larger digital conglomerates. The smaller instances tend to have more distinct and clearly defined expectations, thus facilitating a more cohesive community ethos. Conversely, in expansive platforms with a broad user base, the plurality of values complicates consensus-building and user satisfaction. On the other hand, within smaller instances, moderation may exhibit 'bias', but this is not inherently detrimental. In the context of Mastodon, for instance, such 'bias' may actually contribute to the creation of sanctuaries for specific communities. Administering a Mastodon instance thus requires a nuanced understanding of community expectations, potentially leading to what might be termed 'biased spaces'. This contrasts starkly with centralised social media platforms, where the enormity and diversity of the user base necessitate universal moderation policies, which, in turn, elevate the risk of offensive content due to the sheer volume of interactions. While decentralised platforms offer the potential for safer and more distinctive communities, the complexity of moderation in these contexts underscores the need for a fresh, considered and discerning approach to governance and community engagement.

### 5.3 Power Dynamics and Burdens

Our identification of the intricate power dynamics within Mastodon's administrators and user community resonates with existing literature [114]. Concerns are raised about the potential abuse

of power by Mastodon admins, who could wield significant control, risking detrimental moderation decisions that affect their community members and beyond. At the same time, participant admins highlighted a mutual desire for support and trust between admins and community members. Administrators express a need for enhanced support mechanisms to effectively maintain the health of their communities, establishing a robust system that not only addresses administrative challenges but also fosters a cooperative and trusting environment and power dynamic. In fact, in our study, participants have highlighted approaches they have taken to avoid unintentional authoritarian control by engaging community members in decision-making. This involves selecting trusted community members for the moderation team and encouraging user reports. Some admins also engage with external allies, suggesting that a collaborative moderation approach could alleviate the power imbalances in decentralised social media.

We argue that supporting these power dynamics should be carefully incorporated into any support mechanism designs. While power dynamics has been extensively discussed in existing social media literature [53], it has not always been considered in designing technical tools for the community. For example, supporting content moderation on Mastodon should not be treated as merely addressing administrative needs, but rather as doing so in a way that aligns with the cooperative spirit of decentralised networks. Consequently, content moderation decisions should not be made in isolation but by a representative sub-community that authentically reflects the shared values of the majority. Such approaches would ensure that the decision-making process is inclusive, fair, and in line with the collective ethos of the Mastodon community. Avoiding the creation of a power concentration is essential for maintaining the decentralised and collaborative nature of Mastodon as a whole ecosystem, but may not homogeneously apply to each individual instance. It is beneficial that future designs could be mindful about (mechanisms to support) balancing between empowering administrators and distributing influence across the community. Considering the varied needs and expectations of administrators and community members will ensure that the platform evolves in a way that is inclusive, democratic, and aligned with the principles of decentralisation.

Our study shows that Mastodon admins face burdens including the intricate task of moderating digital communities without resorting to authoritarianism, aiming instead to enhance safety through a calibrated mix of proactive and reactive strategies. This dilemma between community governance and power centralisation has also been shown in other research [35]. Addressing this issue can be challenging. While Dantec and DiSalvo's research [29] indicates that participatory design may aid in this process, though identifying the optimal approach still necessitates further investigation. Nevertheless, the perception of authoritarianism on a few Mastodon instances often emerges as an unintended consequence of proactive moderation strategies aimed at enhancing community safety. This balancing act involves a constant learning process, as they navigate the nuances of maintaining both community standards and the safety of their members, often without formal training or prior experience. The endeavour is not to assert control but to protect, especially within vulnerable communities, and to tailor governance mechanisms that foster safety and inclusivity. In addition, our findings highlight the administrative challenges in managing Mastodon as a networked system [70, 94, 119]. While Mastodon's decentralised structure allows individual communities to establish their own rules and norms, this autonomy can lead to conflicts when interacting with other instances with different policies. Balancing moderation across these diverse environments often requires a nuanced approach to ensure openness and safety without isolating community members. The challenge is exacerbated by instances at different stages of community development and with varying content moderation policies, complicating the task of maintaining a coherent network-wide standard while respecting local norms.

#### 5.4 Design Implications for Decentralised Social Media

The multifaceted nature of the Mastodon platform, with its array of instances catering to a variety of communities, underscores the inherent heterogeneity of society. This decentralised network of social media mirrors the multifariousness of societal values, demonstrating the realism with which digital social structures can reflect broader social dynamics. Each instance on Mastodon has the autonomy to implement moderation policies that may diverge significantly from one another, yet these policies are tailored to meet the specific needs and values of their respective communities. The divergence in standards across Mastodon instances is emblematic of the diversity of human values within society.

Our research indicates that Mastodon instances typically exhibit a pronounced sense of community, each characterised by distinct motivations, challenges, needs, and approaches. This diversity extends to various aspects: safety (with some communities prioritising harassment prevention and others emphasising privacy), content discovery (preferences range from strictly algorithm-free environments to transparent, algorithm-assisted experiences), community growth (differing attitudes towards expansion), and content moderation (varying from proactive safety measures to reactive strategies for greater openness).

This heterogeneity underscores the inadequacy of a universal solution across Mastodon instances. In contrast to traditional social media platforms governed by a singular entity with unified objectives and regulations, decentralised social media thrives on diversity and customisation. Consequently, the design of tools and mechanisms for such platforms must account for the unique values and needs of specific communities. For instance, the development of algorithms demands an in-depth comprehension of the community's stance towards such technology, their values, and concerns. Similarly, implementing a generic, pre-trained AI detection tool for moderation may not be effective in a decentralised context. Instead, it requires a nuanced understanding of the target community, considering factors like the instance's specific moderation policies, definitions of harmful content, and preferred moderation styles. In this context, prior works in CSCW on the design of personalised content moderation tools [62, 67, 99] could serve as valuable start points, paving the way for the necessity of building tailored approaches attuned to the distinct values and needs of each community. Universal, one-size-fits-all tools can be unsuitable for this varied landscape, highlighting the need for designs that are as diverse as the communities they serve.

This research primarily focuses on examining decentralised social media; however, it also provides insights applicable to a broader spectrum of social computing systems, including centralised social media platforms. Some of our findings resonate with previous studies on centralised community-driven platforms like Reddit, echoing with Hwang and Foote's observations on the appeal of smaller communities [57] and offering empirical explanations for Nicholson et al.'s observations regarding the prevalent rules against harassment and hate issues on Mastodon compared to Reddit [84]. These rules contribute to Mastodon as a 'safe space' for diverse identities. Our research further unpacks the similarities and differences between Mastodon and Reddit, and provides building blocks for understanding the dynamics of user interaction and governance across different social media platforms. Furthermore, our research elucidates the motivations for transitioning from centralised social media platforms to Mastodon, highlight the potential deficiencies of centralised platforms. Our findings suggest that improvements in safety measures for diverse communities including marginalised groups, along with increased user control over content feeds and enhanced data autonomy, could potentially enhance centralised social media platforms.

## 6 Limitations and Future Work

This research has several limitations, particularly concerning the size and composition of the study population. The participants signed up for our study may already have heightened awareness about issues like citizen empowerment. Another limitation of our approach lies in the fact that all data collected is self-reported. This means the insights are confined to what the participants deemed significant or relevant, based on their personal perspectives. Consequently, our study's findings might predominantly mirror the perceptions and experiences of general demographics or LGBTQ+ groups, and more specifically, those of administrators, as opposed to moderators or regular users. Nevertheless, we believe that the work presented here paves the way for future ethnographic observations. These could yield a more direct comprehension of the practices and further contribute to critical insights on the future of decentralised social media. Finally, decentralised social media encompasses a wide range of platforms, each with its unique set of challenges and considerations. Our decision to choose Mastodon as our example platform is based on its significance and prominence in the decentralised space, which may influence our findings, particularly in terms of issues that are most relevant to this specific platform. Future research will focus on developing methods that enable users on these platforms to enjoy experiences more closely aligned with their values and the specific needs of their communities, through methods like co-design workshops. These workshops will explore various models of content moderation and community governance designs, and investigate strategies to empower users more effectively on these platforms.

## 7 Conclusion

In the evolving social media landscape, decentralised platforms like Mastodon, an open-source, federated platform, have become notable alternatives to traditional centralised networks, offering user autonomy and community governance. This paper presents a systematic examination of the experiences of Mastodon administrators, offering vital insights into the dynamics of decentralised social media platforms. The research, grounded in semi-structured interviews with 16 Mastodon administrators, reveals the motivations, challenges, and strategies employed by those at the forefront of this emerging landscape. Key findings indicate that while decentralisation offers enhanced user control, privacy, and the creation of safe, inclusive environments, it also brings significant challenges such as content moderation and user adoption. The study highlights the balance between the benefits of a citizen-powered, customized social media experience and the increased responsibilities of platform governance. This exploration is crucial for understanding the evolving trajectory of decentralised social media, paving the way towards a more citizen-powered future.

## Acknowledgments

This work was supported by the UK Engineering and Physical Sciences Research Council [1535259] and the Ethical Web and Data Architecture programme of the Oxford Martin School.

## References

- [1] [n. d.]. <https://the-federation.info/>
- [2] [n. d.]. The Diaspora\* Project. <https://diasporafoundation.org/>.
- [3] [n. d.]. Hubzilla. <https://hubzilla.org/page/info/discover>.
- [4] [n. d.]. Mastodon Server Covenant. <https://joinmastodon.org/covenant>
- [5] Roel Roscam Abbing, Cade Diehm, and Shahed Warreth. 2023. Decentralised social media. *Internet Policy Review* 12, 1 (2023).
- [6] Yong-Yeol Ahn, Seungyeop Han, Haewoon Kwak, Sue Moon, and Hawoong Jeong. 2007. Analysis of topological characteristics of huge online social networking services. In *Proceedings of the 16th International Conference on World Wide Web*. 835–844.

- [7] Tawfiq Ammari, Momina Nofal, Mustafa Naseem, and Maryam Mustafa. 2022. Moderation as Empowerment: Creating and Managing Women-Only Digital Safe Spaces. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 313 (nov 2022), 36 pages. <https://doi.org/10.1145/3555204>
- [8] Ishaku Hassan Anaobi, Aravindh Raman, Ignacio Castro, Haris Bin Zia, Damilola Ibosiola, and Gareth Tyson. 2023. Will Admins Cope? Decentralized Moderation in the Fediverse. In *Proceedings of the ACM Web Conference 2023*. 3109–3120.
- [9] Ricardo Baeza-Yates. 2018. Bias on the web. *Commun. ACM* 61, 6 (2018), 54–61.
- [10] Chelsea Barabas, Neha Narula, and Ethan Zuckerman. 2017. Defending Internet freedom through decentralization: back to the future. *The Center for Civic Media & The Digital Currency Initiative MIT Media Lab* (2017).
- [11] Kathy Baxter, Catherine Courage, and Kelly Caine. 2015. *Understanding your users: a practical guide to user research methods*. Morgan Kaufmann.
- [12] Elysse Bell. 2023. How mastodon makes money. <https://www.investopedia.com/how-mastodon-makes-money-7482865>
- [13] Sophie Bishop. 2021. Influencer management tools: Algorithmic cultures, brand safety, and bias. *Social media+ society* 7, 1 (2021), 20563051211003066.
- [14] Lindsay Blackwell, Jean Hardy, Tawfiq Ammari, Tiffany Veinot, Cliff Lampe, and Sarita Schoenebeck. 2016. LGBT parents and social media: Advocacy, privacy, and disclosure during shifting social movements. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 610–622.
- [15] Oscar Borgogno and Giuseppe Colangelo. 2019. Data sharing and interoperability: Fostering innovation and competition through APIs. *Computer Law & Security Review* 35, 5 (2019), 105314.
- [16] Richard E Boyatzis. 1998. *Transforming qualitative information: Thematic analysis and code development*. sage.
- [17] Joshua Braun. 2023. Journalism, Media Research, and Mastodon: Notes on the Future. *Digital Journalism* (2023), 1–8.
- [18] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [19] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and psychotherapy research* 21, 1 (2021), 37–47.
- [20] Allison J Brown. 2020. “Should I stay or should I leave?”: Exploring (dis) continued Facebook use after the Cambridge Analytica scandal. *Social media+ society* 6, 1 (2020), 2056305120913884.
- [21] Amy S Bruckman. 2022. *Should you believe Wikipedia?: online communities and the construction of knowledge*. Cambridge University Press.
- [22] Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, and Krishna Gummadi. 2010. Measuring user influence in twitter: The million follower fallacy. In *Proceedings of the international AAAI conference on web and social media*, Vol. 4. 10–17.
- [23] Eshwar Chandrasekharan, Mattia Samory, Shagun Jhaver, Hunter Charvat, Amy Bruckman, Cliff Lampe, Jacob Eisenstein, and Eric Gilbert. 2018. The Internet’s hidden rules: An empirical study of Reddit norm violations at micro, meso, and macro scales. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–25.
- [24] Yan Chen, Jack I Richter, and Pankaj C Patel. 2021. Decentralized governance of digital platforms. *Journal of Management* 47, 5 (2021), 1305–1337.
- [25] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences* 118, 9 (2021), e2023301118.
- [26] Rosemary Clark-Parsons. 2018. Building a digital Girl Army: The cultivation of feminist safe spaces online. *New Media & Society* 20, 6 (2018), 2125–2144.
- [27] James Clayton and Peter Hoskins. 2022. Elon Musk takes control of Twitter in \$44bn deal. <https://www.bbc.co.uk/news/technology-63402338>
- [28] Kimberle Crenshaw. 1991. Women of color at the center: Selections from the third national conference on women of color and the law: Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review* 43, 6 (1991), 1241–1279.
- [29] Christopher A Le Dantec and Carl DiSalvo. 2013. Infrastructuring and the formation of publics in participatory design. *Social Studies of Science* 43, 2 (2013), 241–264.
- [30] A Datta, S Buchegger, L-H Vu, T Strufe, and K Rzađca. 2010. Decentralized online social networks. In *Handbook of social network technologies and applications*, B Furht (Ed.). Springer, 349–378.
- [31] Ronald J Deibert. 2019. Three painful truths about social media. *J. Democracy* 30 (2019), 25.
- [32] Michael Ann DeVito, Ashley Marie Walker, and Julia R. Fernandez. 2021. Values (Mis)Alignment: Exploring Tensions Between Platform and LGBTQ+ Community Design Values. 5, CSCW1, Article 88 (apr 2021), 27 pages. <https://doi.org/10.1145/3449162>



- [33] Trinh Viet Doan, Tat Dat Pham, Markus Oberprieler, and Vaibhav Bajpai. 2020. Measuring decentralized video streaming: A case study of dtube. In *2020 IFIP Networking Conference (Networking)*. IEEE, 118–126.
- [34] Kevin Driscoll. 2022. *The modem world: A prehistory of social media*. Yale University Press.
- [35] Brianna Dym, Namita Pasupuleti, and Casey Fiesler. 2022. Building a pillowfort: Political tensions in platform design and policy. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–23.
- [36] KJ Feng, Xander Koo, Lawrence Tan, Amy Bruckman, David W McDonald, and Amy X Zhang. 2024. Mapping the Design Space of Teachable Social Media Feed Experiences. *arXiv preprint arXiv:2401.14000* (2024).
- [37] Casey Fiesler, Jialun Jiang, Joshua McCann, Kyle Frye, and Jed Brubaker. 2018. Reddit rules! characterizing an ecosystem of governance. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 12.
- [38] Seth Flaxman, Sharad Goel, and Justin M Rao. 2016. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly* 80, S1 (2016), 298–320.
- [39] Christian Fuchs. 2011. Web 2.0, presumption, and surveillance. *Surveillance & Society* 8, 3 (2011), 288–309.
- [40] Christine Geeng and Alexis Hiniker. 2021. LGBTQ privacy concerns on social media. *arXiv preprint arXiv:2112.00107* (2021).
- [41] Robert W Gehl and Diana Zulli. 2022. The digital covenant: non-centralized platform governance on the mastodon social network. *Information, Communication & Society* (2022), 1–17.
- [42] Sarah A Gilbert. 2020. "I run the world's largest historical outreach project and it's on a cesspool of a website." Moderating a Public Scholarship Site on Reddit: A Case Study of r/AskHistorians. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–27.
- [43] Tarleton Gillespie. 2018. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- [44] Robert Gorwa, Reuben Binns, and Christian Katzenbach. 2020. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society* 7, 1 (2020), 2053951719897945.
- [45] B Guidi, M Conti, A Passarella, and L Ricci. 2018. Managing social contents in decentralized online social networks: a survey. *Online Social Networks and Media* 7 (2018), 12–29.
- [46] Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2016. Digital Footprints and Changing Networks During Online Identity Transitions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). Association for Computing Machinery, New York, NY, USA, 2895–2907. <https://doi.org/10.1145/2858036.2858136>
- [47] Oliver L Haimson, Justin Buss, Zu Weinger, Denny L Starks, Dyke Gorrell, and Briar Sweetbriar Baron. 2020. Trans time: Safety, privacy, and content warnings on a transgender-specific social media site. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–27.
- [48] Oliver L Haimson, Daniel Delmonaco, Peipei Nie, and Andrea Wegner. 2021. Disproportionate removals and differing content moderation experiences for conservative, transgender, and black social media users: Marginalization and moderation gray areas. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–35.
- [49] Camille Harris, Amber Gayle Johnson, Sadie Palmer, Diyi Yang, and Amy Bruckman. 2023. "Honestly, I Think TikTok has a Vendetta Against Black Creators": Understanding Black Content Creator Experiences on TikTok. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–31.
- [50] Anaobi Ishaku Hassan, Aravindh Raman, Ignacio Castro, and Gareth Tyson. 2021. The impact of Capitol Hill on Pleroma: the case for decentralised moderation. In *Proceedings of the CoNEXT Student Workshop*. 1–2.
- [51] Anaobi Ishaku Hassan, Aravindh Raman, Ignacio Castro, Haris Bin Zia, Emiliano De Cristofaro, Nishanth Sastry, and Gareth Tyson. 2021. Exploring content moderation in the decentralised web: The pleroma case. In *Proceedings of the 17th International Conference on emerging Networking EXperiments and Technologies*. 328–335.
- [52] Michael Hauben and Ronda Hauben. 1997. *Netizens: On the history and impact of Usenet and the Internet*. IEEE Computer Society Press.
- [53] Anne Helmond. 2015. The platformization of the web: Making web data platform ready. *Social media+ society* 1, 2 (2015), 2056305115603080.
- [54] Bernie Hogan. 2016. Social media giveth, social media taketh away: Facebook, friendships, and APIs. *International Journal of Communication, Forthcoming* (2016).
- [55] Kalley Huang. 2022. What Is Mastodon and Why Are People Leaving Twitter for It? *The New York Times* (2022).
- [56] Lee Humphreys, Phillipa Gill, and Balachander Krishnamurthy. 2010. How much is too much? Privacy issues on Twitter. In *Conference of international communication association, singapore*.
- [57] Sohyeon Hwang and Jeremy D Foote. 2021. Why do people participate in small online communities? *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–25.
- [58] Sohyeon Hwang, Priyanka Nanayakkara, and Yan Shvartzshnaider. 2023. Whose Policy? Privacy Challenges of Decentralized Platforms. In *CHI'23 Workshops: Designing Technology and Policy Simultaneously: Towards A Research Agenda and New Practice*.

- [59] Jack Jamieson and Naomi Yamashita. 2023. Escaping the Walled Garden? User Perspectives of Control in Data Portability for Social Media. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–27.
- [60] Ujun Jeong, Paras Sheth, Anique Tahir, Faisal Alatawi, H Russell Bernard, and Huan Liu. 2023. Exploring platform migration patterns between twitter and mastodon: A user behavior study. *arXiv preprint arXiv:2305.09196* (2023).
- [61] Shagun Jhaver, Seth Frey, and Amy X Zhang. 2023. Decentralizing Platform Power: A Design Space of Multi-level Governance in Online Social Platforms. *Social Media+ Society* 9, 4 (2023), 20563051231207857.
- [62] Shagun Jhaver, Alice Qian Zhang, Quan Ze Chen, Nikhila Natarajan, Ruotong Wang, and Amy X Zhang. 2023. Personalizing content moderation on social media: User perspectives on moderation choices, interface design, and labor. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–33.
- [63] Anna-Katharina Jung, Stefan Stieglitz, Tobias Kissmer, Milad Mirbabaie, and Tobias Kroll. 2022. Click me...! The influence of clickbait on user engagement in social media and the role of digital nudging. *Plos one* 17, 6 (2022), e0266743.
- [64] Moira Kenney. 2001. *Mapping gay LA: The intersection of place and politics*. Temple University Press.
- [65] Brent Kitchens, Steven L Johnson, and Peter Gray. 2020. Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption. *MIS quarterly* 44, 4 (2020).
- [66] Iga Kozłowska. 2018. Facebook and data privacy in the age of Cambridge Analytica. *The Henry M. Jackson School of International Studies* (2018), 1.
- [67] Tina Kuo, Alicia Hernani, and Jens Grossklags. 2023. The Unsung Heroes of Facebook Groups Moderation: A Case Study of Moderation Practices and Tools. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–38.
- [68] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web*. 591–600.
- [69] Michael Kwet. 2020. Fixing social media: toward a democratic digital commons. *Markets, Globalization & Development Review* 5, 1 (2020).
- [70] Lucio La Cava, Sergio Greco, and Andrea Tagarelli. 2021. Understanding the growth of the Fediverse through the lens of Mastodon. *Applied network science* 6 (2021), 1–35.
- [71] Leanna Lucero. 2017. Safe spaces in online places: Social media and LGBTQ youth. *Multicultural Education Review* 9, 2 (2017), 117–128.
- [72] Christopher Lueg and Danyel Fisher. 2012. *From Usenet to CoWebs: interacting with social information spaces*. Springer Science & Business Media.
- [73] Jens-Erik Mai. 2016. Big data privacy: The datafication of personal information. *The Information Society* 32, 3 (2016), 192–199.
- [74] Kate Mannell and Eden T Smith. 2022. Alternative social media and the complexities of a more participatory culture: A view from Scuttlebutt. *Social Media+ Society* 8, 3 (2022), 20563051221122448.
- [75] Huina Mao, Xin Shuai, and Apu Kapadia. 2011. Loose tweets: an analysis of privacy leaks on twitter. In *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society*. 1–12.
- [76] Francisco J Martínez-López, Yangchun Li, and Susan M Young. 2022. Social Media Monetization and Demonetization: Risks, Challenges, and Potential Solutions. In *Social Media Monetization: Platforms, Strategic Models and Critical Success Factors*. Springer, 185–214.
- [77] Mastodon. 2016. <https://joinmastodon.org/>
- [78] Mastodon. 2024. What is Mastodon? <https://docs.joinmastodon.org/>
- [79] Ariadna Matamoros-Fernández and Johan Farkas. 2021. Racism, hate speech, and social media: A systematic review and critique. *Television & New Media* 22, 2 (2021), 205–224.
- [80] Ulises A Mejias and Nick Couldry. 2019. Datafication. *Internet Policy Review* 8, 4 (2019).
- [81] Filippo Menczer, David Crandall, Yong-Yeol Ahn, and Apu Kapadia. 2023. Addressing the harms of AI-generated inauthentic content. *Nature Machine Intelligence* (2023), 1–2.
- [82] Michela Moscufo and Mike Dobuski. 2022. LGBTQ+ community facing increased social media bias, author says. <https://abcnews.go.com/US/lgbtq-community-facing-increased-social-media-bias-author/story?id=85463533>
- [83] Seth A Myers, Aneesh Sharma, Pankaj Gupta, and Jimmy Lin. 2014. Information network or social network? The structure of the Twitter follow graph. In *Proceedings of the 23rd International Conference on World Wide Web*. 493–498.
- [84] Matthew N Nicholson, Brian C Keegan, and Casey Fiesler. 2023. Mastodon Rules: Characterizing Formal Rules on Popular Mastodon Instances. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*. 86–90.
- [85] Fayika Farhat Nova, Michael Ann DeVito, Pratyasha Saha, Kazi Shohanur Rashid, Shashwata Roy Turzo, Sadia Afrin, and Shion Guha. 2021. "Facebook Promotes More Harassment": Social Media Ecosystem, Skill and Marginalized Hijra Identity in Bangladesh. 5, CSCW1, Article 157 (apr 2021), 35 pages. <https://doi.org/10.1145/3449231>

- [86] Elias Capello Oliver L. Haimson, Avery Dame-Griff and Zahari Richter. 2021. Tumblr was a trans technology: the meaning, importance, history, and future of trans technologies. *Feminist Media Studies* 21, 3 (2021), 345–361. <https://doi.org/10.1080/14680777.2019.1678505> arXiv:<https://doi.org/10.1080/14680777.2019.1678505>
- [87] Lauren Olson, Emitzá Guzmán, and Florian Kunneman. 2023. Along the Margins: Marginalized Communities' Ethical Concerns about Social Platforms. *arXiv preprint arXiv:2304.08882* (2023).
- [88] Jose Ortiz, Amber Young, Michael Myers, Rudolph T Bedeley, Donal Carbaugh, Hameed Chughtai, Elizabeth Davidson, Jordana George, Janis Gogan, Steven Gordon, et al. 2019. Giving voice to the voiceless: The use of digital technologies by marginalized groups. (2019).
- [89] Katy E Pearce, Amy Gonzales, and Brooke Foucault Welles. 2020. Introduction: marginality and social media. *Social media+ society* 6, 3 (2020), 2056305120930413.
- [90] PeerTube. 2018. <https://joinpeertube.org>
- [91] Bryan Pfaffenberger. 1996. "If I want it, it's OK": Usenet and the (Outer) limits of free speech. *The Information Society* 12, 4 (1996), 365–386.
- [92] Anthony T. Pinter, Morgan Klaus Scheuerman, and Jed R. Brubaker. 2021. Entering Doors, Evading Traps: Benefits and Risks of Visibility During Transgender Coming Outs. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 272 (Jan 2021), 27 pages. <https://doi.org/10.1145/3434181>
- [93] Pleroma. 2018. <https://pleroma.social/>
- [94] 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*. 217–229.
- [95] Mark van Rijmenam. 2023. Privacy in the age of AI: Risks, challenges and solutions. <https://www.thedigitalspeaker.com/privacy-age-ai-risks-challenges-solutions/>
- [96] Amy Ross Arguedas, C Robertson, Richard Fletcher, and R Nielsen. 2022. Echo chambers, filter bubbles, and polarisation: A literature review. (2022).
- [97] Alan Z Rozenshtein. 2022. Moderating the Fediverse: Content Moderation on Distributed Social Media. (2022).
- [98] Siavosh Sahebi and Paul Formosa. 2022. Social media and its negative impacts on autonomy. *Philosophy & Technology* 35, 3 (2022), 70.
- [99] Julia Sasse and Jens Grossklags. 2023. Breaking the Silence: Investigating Which Types of Moderation Reduce Negative Effects of Sexist Social Media Content. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–26.
- [100] Florian Saurwein and Charlotte Spencer-Smith. 2021. Automated trouble: The role of algorithmic selection in harms on social media platforms. *Media and Communication* 9, 4 (2021), 222–233.
- [101] Morgan Klaus Scheuerman, Stacy M Branham, and Foad Hamidi. 2018. Safe spaces and safe places: Unpacking technology-mediated experiences of safety and harm with transgender people. *Proceedings of the ACM on Human-computer Interaction* 2, CSCW (2018), 1–27.
- [102] Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter. 2017. Intersectional HCI: Engaging Identity through Gender, Race, and Class (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 5412–5427. <https://doi.org/10.1145/3025453.3025766>
- [103] Jonathan E Schroeder. 2021. Reinscribing gender: social media, algorithms, bias. *Journal of marketing management* 37, 3–4 (2021), 376–378.
- [104] Nick Srnicek. 2017. *Platform capitalism*. John Wiley & Sons.
- [105] Hannah Tam, Karthik S. Bhat, Priyanka Mohindra, and Neha Kumar. 2023. Learning to Navigate Health Taboos through Online Safe Spaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (, Hamburg, Germany,) (*CHI '23*). Association for Computing Machinery, New York, NY, USA, Article 295, 15 pages. <https://doi.org/10.1145/3544548.3580708>
- [106] Omer Tene and Jules Polonetsky. 2012. Big data for all: Privacy and user control in the age of analytics. *Nw. J. Tech. & Intell. Prop.* 11 (2012), 239.
- [107] David Thiel and Renee DiResta. 2023. Addressing Child Exploitation on Federated Social Media. <https://cyber.fsi.stanford.edu/io/news/addressing-child-exploitation-federated-social-media>
- [108] Amanda L Traud, Peter J Mucha, and Mason A Porter. 2012. Social structure of facebook networks. *Physica A: Statistical Mechanics and its Applications* 391, 16 (2012), 4165–4180.
- [109] Carmela Troncoso, Marios Isaakidis, George Danezis, and Harry Halpin. 2017. Systematizing decentralization and privacy: Lessons from 15 years of research and deployments. *arXiv preprint arXiv:1704.08065* (2017).
- [110] Emily Van Duyn. 2020. Mainstream marginalization: Secret political organizing through social media. *Social Media+ Society* 6, 4 (2020), 2056305120981044.
- [111] Emily A Vogels. 2021. The state of online harassment. *Pew Research Center* 13 (2021), 625.
- [112] Stanislav Vojří, Zdeněk Smutný, and Jan Kučera. 2020. Social and Technical Aspects of Re-decentralized web. *IDIMT-2020 Digitalized Economy, Society and Information Management* (2020), 107–116.

- [113] W3C. 2018. ActivityPub. <https://www.w3.org/TR/activitypub/>.
- [114] Ashley Marie Walker and Michael A. DeVito. 2020. "'More Gay' Fits in Better": Intracommunity Power Dynamics and Harms in Online LGBTQ+ Spaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (, Honolulu, HI, USA,) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3313831.3376497>
- [115] Ge Wang, Jun Zhao, Samantha-Kaye Johnston, Zhilin Zhang, Max Van Kleek, and Nigel Shadbolt. 2024. CHAITok: A Proof-of-Concept System Supporting Children's Sense of Data Autonomy on Social Media. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.
- [116] Qunfang Wu and Bryan Semaan. 2023. "How Do You Quantify How Racist Something Is?": Color-Blind Moderation in Decentralized Governance. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–27.
- [117] Fouzia Younas, Mustafa Naseem, and Maryam Mustafa. 2020. Patriarchy and Social Media: Women Only Facebook Groups as Safe Spaces for Support Seeking in Pakistan. In *Proceedings of the 2020 International Conference on Information and Communication Technologies and Development* (Guayaquil, Ecuador) (*ICTD '20*). Association for Computing Machinery, New York, NY, USA, Article 11, 11 pages. <https://doi.org/10.1145/3392561.3394639>
- [118] Haris Bin Zia, Jiahui He, Aravindh Raman, Ignacio Castro, Nishanth Sastry, and Gareth Tyson. 2023. Flocking to mastodon: Tracking the great twitter migration. *arXiv preprint arXiv:2302.14294* (2023).
- [119] Matteo Zignani, Sabrina Galto, and Gian Paolo Rossi. 2018. Follow the "Mastodon": Structure and evolution of a decentralized online social media. In *ICWSM*. 541–550.
- [120] Matteo Zignani, Christian Quadri, Sabrina Gaito, Hocine Cherifi, and Gian Paolo Rossi. 2019. The Footprints of a Mastodon: How a Decentralized Architecture Influences Online Social Relationships. In *INFOCOM 2019*. 472–477.
- [121] Matteo Zignani, Christian Quadri, Alessia Galdeman, Sabrina Gaito, and Gian Paolo Rossi. 2019. Mastodon content warnings: Inappropriate contents in a microblogging platform. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 639–645.
- [122] Franziska Zimmer, Katrin Scheibe, Mechtilde Stock, and Wolfgang G Stock. 2019. Echo chambers and filter bubbles of fake news in social media. Man-made or produced by algorithms. In *8th annual arts, humanities, social sciences & education conference*. 1–22.
- [123] Shoshana Zuboff. 2015. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of information technology* 30, 1 (2015), 75–89.
- [124] Shoshana Zuboff. 2019. *The age of surveillance capitalism: The fight for a human future at the new frontier of power: Barack Obama's books of 2019*. Profile books.
- [125] Diana Zulli, Miao Liu, and Robert Gehl. 2020. Rethinking the "social" in "social media": Insights into topology, abstraction, and scale on the Mastodon social network. *New Media & Society* 22, 7 (2020), 1188–1205.
- [126] Oana Ștefăniță and Diana-Maria Buf. 2021. Hate speech in social media and its effects on the LGBT community: A review of the current research. *Romanian Journal of Communication and Public Relations* 23, 1 (2021), 47–55.

Received January 2024; revised April 2024; accepted May 2024