ISD | Powering solutions to extremism, hate and disinformation

CASM technology

Researching the Evolving Online Ecosystem: Telegram, Discord and Odysee

Henry Tuck, Jakob Guhl, Julia Smirnova, Lea Gerster and Oliver Marsh

WARNING: This report contains Figures and descriptions of posts that readers may find extremely upsetting. These include calls for violence, the glorification of terrorism and the dehumanisation of minority communities.

About this publication

Harmful actors use an ever-expanding range of digital spaces to spread harmful ideologies and undermine human rights and democracy online. Understanding their evolving ideas, online networks and activities is critical to developing a more comprehensive evidence base to inform effective and proportional efforts to counter them. But creating that evidence base can challenge the technical capabilities, resources, and even ethical and legal boundaries of research. We are concerned that all these may be getting worse, just as the options for spreading harm online increase. It should therefore be of concern that in many instances it is increasingly hard to conduct digital research systematically, ethically and legally. This results in a situation where difficult trade-offs have to be made between competing goods, including the desire to understand and mitigate harmful content and behaviour online, the preservation of privacy and the adherence to legal agreements. We argue in this report that this does not need to be the case; solutions are available, and actions should be taken as soon as possible to ensure a future-proof scenario in which researchers have the tools to monitor, track and analyse harmful content and behaviour systematically, ethically and legally. This report outlines the findings from the research phase of a project by the Institute for Strategic Dialogue (ISD) and CASM Technology, and funded by Omidyar Network. The aim of the project is to identify and test research methodologies to monitor and analyse small, closed or hardly moderated platforms. It provides applied examples and evidence for the limitations and dilemmas encountered by researchers. In three small research case studies, focusing on Telegram, Discord and Odysee in German, English and French respectively, we seek to apply different methodological approaches to analyse platforms that primarily present technological, ethical and legal, or fragmentation barriers.

About the Authors

Jakob Guhl is a Senior Research Manager at ISD, where he works within the Digital Research Unit and with ISD Germany. His research focuses on the far-right, Islamist extremism, hate speech, disinformation and conspiracy theories. Jakob has been invited to present his research to the German Ministry of the Justice and provided evidence to the German Minister of the Interior on how to strengthen prevention against right-wing extremism and antisemitism. Oliver Marsh is the founder of The Data Skills Consultancy, which supports work at the intersection of data skills and soft skills. Previously as a government official, he helped create the Rapid Response Unit in Downing Street and the UK's post-Brexit Data Adequacy capability in DCMS. He is a Fellow of the think-tank Demos, a Policy Fellow of the Royal Academy of Engineering, and an Honorary Research Associate of the Science and Technology Studies Department at UCL. Henry Tuck is the Head of Digital Policy at ISD, where he leads work on digital regulation and tech company responses to terrorism, extremism, hate and dis/ misinformation online. Henry oversees ISD's Digital Policy Lab (DPL) programme and advisory work on key digital regulation proposals in Europe and Five Eyes countries, and collaborates with ISD's Digital Analysis Unit to translate research into actionable digital policy recommendations.

Lea Gerster is an Analyst at ISD and ISD Germany. She works across a range of projects focused on the proliferation of extremist ideologies and disinformation in the English- and German-speaking part of the internet. Previously, Lea worked for two years in online extremism-related roles at TRD Policy and the Centre on Radicalisation and Terrorism. She also worked as an intern for the Swiss foreign office and as a volunteer at a Japanese tea farm. She is the co-author of the ISD report Crisis and Loss of Control: German-Language Digital Extremism in the Context of the COVID-19 Pandemic. She holds an MA War Studies from King's College London and a BA in history, Japan studies, and political sciences.

Julia Smirnova is a Senior Analyst at ISD and ISD Germany, focusing on proliferation of disinformation, conspiracy myths, hate speech and extremist ideologies online. She is co-author of ISD reports Digitale Gewalt und Desinformation gegen Spitzenkandidat:innen vor der Bundestagswahl 2021, Ein Virus des Misstrauens: Der russische Staatssender RT DE und die deutsche Corona-Leugner-Szene and Desinformationskampagnen gegen die Wahl: Befunde aus Sachsen-Anhalt. Before joining ISD, Julia was working as a journalist writing for Der Spiegel and Die Zeit among others. As the Moscow correspondent for Die Welt, she was reporting on Russian politics, the annexation of Crimea and the war in Eastern Ukraine. She holds an MA in Conflict, Security and Development from King's College London.

Oliver Marsh is a consultant at CASM Technology and the founder of The Data Skills Consultancy, which supports work at the intersection of data skills and soft skills. Previously, as a government official, he helped create the Rapid Response Unit in Downing Street and the UK's post-Brexit Data Adequacy capability in DCMS. He is a Fellow of the think tank Demos, a Policy Fellow of the Royal Academy of Engineering, and an Honorary Research Associate of the Science and Technology Studies Department at UCL.

Henry Tuck is the Head of Digital Policy at ISD, where he leads work on digital regulation and tech company responses to terrorism, extremism, hate and dis/misinformation online. Henry oversees ISD's Digital Policy Lab (DPL) programme and advisory work on key digital regulation proposals in Europe and Five Eyes countries, and collaborates with ISD's Digital Analysis Unit to translate research into actionable digital policy recommendations.

Acknowledgments

This report would not have been possible without funding support from Omidyar Network. We would like to express our gratitude to Wafa Ben-Hassine, Anamitra Deb and Emma Leiken for their vision, continuing support and insightful feedback.

The authors would also like thank Nestor Prieto Chavana, Jack Pay and Carl Miller at CASM, whose technical expertise and advice has been vital to this research.

ISD | Powering solutions to extremism, hate and disinformation

Amman | Berlin | London | Paris | Washington DC

Contents

Executive Summary **5**
Glossary **8**
Introduction **10**
Case Study 1: Telegram **12**
– Platform Overview 13
– Researching Telegram 15
– Analysis of Telegram Communities: Key Findings 21
Case Study 2: Discord **26**
– Platform Overview 27
– Researching Discord 28
– Analysis of Discord Communities: Key Findings 32
Case Study 3: Odysee **38**
– Platform Overview 39
– Researching Odysee 41
– Analysis of Odysee Communities: Key Findings 43
Conclusions and Recommendations **50**
– Implications for Researchers 50
– Policy Context 51
– Recommendations 55
Endnotes **58**

Executive Summary

Harmful actors use an ever-expanding range of digital spaces to spread harmful ideologies and undermine human rights and democracy online. Understanding their evolving ideas, online networks and activities is critical to the development of a more comprehensive evidence base to inform effective and proportional efforts to counter them. But generating that evidence base can challenge the technical capabilities, resources and even ethical and legal boundaries of research. We are concerned that these issues may be worsening just as the options for spreading harm online increase.

This difficulty in conducting digital research systematically, ethically and legally results in a situation where trade-offs have to be made between competing priorities, including the desire to understand and mitigate harmful content and behaviours online, the preservation of privacy and the adherence to legal agreements. We argue in this report that this does not need to be the case; solutions are available, and actions should be taken as soon as possible to ensure a future-proof scenario in which researchers have the tools to monitor, track and analyse harmful content and behaviours systematically, ethically and legally.

This report outlines the findings from the research phase of a project by the Institute for Strategic Dialogue (ISD) and CASM Technology; it is funded by Omidyar Network. The aim of the project is to identify and test research methodologies for monitoring and analysing small, closed or hardly moderated platforms. The report provides applied examples and evidence for the limitations and dilemmas encountered by researchers. In three short research case studies, focusing on Telegram, Discord and Odysee (in German, English and French respectively), we seek to apply different methodological approaches to analyse platforms that primarily present technological, ethical and legal, or fragmentation barriers.

Key Findings: Barriers to Research

- **Overall, it is possible but difficult to conduct digital research on platforms like Telegram, Discord and Odysee in a way that is simultaneously systematic, ethical and legal.** Ethical considerations around the right to privacy or expectations of privacy prevented us from researching more harmful Telegram communities, while legal concerns around breaking contract law stopped us from systematic data collection of Discord servers. Ethical and legal barriers therefore limit researchers' ability to systematically study harmful communities on these platforms. Although there were fewer ethical concerns with studying Odysee, given its public nature, a combination of technical complexity and unclear Terms of Service (TOS) regarding data collection also made more comprehensive analysis challenging.
- **On Telegram, platform functionalities enable the admins of arguably public spaces to prevent systematic research of harmful communities.** Telegram allows users to create 'private' groups and channels, which cannot be systematically found and researched. As these channels may have thousands of members, they are straining the definition of private. Joining these channels may also require researchers to be deceitful about their identity, presenting ethical barriers.
- **On Discord, fragmentation, ethical and legal barriers combined to limit the scope of our research findings.** Discord's Application Programming Interface (API) and non-transparent search functions provided by the platform, as well as other third-party software present fragmentation barriers. At the same time, Discord's TOS prohibit the collection of user data via its API, likely posing legal risks for researchers. Similarly, the questionnaires to enter groups often required significant deception or the promotion of problematic beliefs, which prevented us from joining many servers of interest.
- **Odysee's blockchain-based design (atechnological barrier) presented fewer insurmountable barriers to research than might be expected.** Transactions using the LBC cryptocurrency are in fact publicly visible and provide additional data points. However, this requires technical expertise and substantial effort to investigate and amalgamate different tools; the novelty of blockchain means that many organisations may not have access to such expertise and fewer open-source tools may be available. The audiovisual nature of the platform presented further challenges as researchers needed to manually inspect content. Vague TOS on Odysee also caused uncertainty around what is appropriate or legally permissible.

- **Technological, ethical and legal, and fragmentation barriers on Telegram, Discord and Odysee are therefore highly interconnected.** For example, legal considerations prevented us from overcoming fragmentation barriers through systematic search methods on Discord. Similarly, technological features on Telegram created ethical barriers.

Key Findings: Implications for Researchers

Our research during this project has demonstrated how platforms can, whether deliberately or inadvertently, inhibit public interest research in several ways.

- **Technological choices by platforms can prevent access to data or hosted content or make data collection unnecessarily burdensome from a technical perspective,** for example, where platforms' front-end or back-end architectures are fragmented or poorly maintained, or the tools available to access data are poorly documented. The expertise required to explore the technical possibilities across a variety of different platforms may not be available to all researchers, particularly in civil society.
- **Unclear platform legal terms can create additional burdens or risks for researchers,** raising the costs and barriers to entry for online research. We were fortunate to have access to external, pro bono legal support that enabled us to more carefully assess platforms' TOS and any resulting legal risks related to data access.
- **Platforms can deliberately invoke legal or ethical concerns to justify restricting researchers' access to data,** including around data protection or users' rights to privacy. This can occur even in cases where users have actively consented to the sharing and use of their data for specific, limited research purposes, or where the designation of online spaces as private can create ethical uncertainty despite their being easily and widely accessible to both users and researchers in practice.
- **There is a danger that the research community's scarce resources for addressing these challenges are being employed in a disconnected and siloed way.** Consequently, without strong mechanisms and incentives for greater collaboration, opportunities to achieve greater economies of scale are being missed
- **Barriers to online research can create a problematic set of incentives for researchers** when it comes to balancing or mitigating different legal and ethical risks. For example, deception may actually serve to reduce legal risks for researchers utilising third-party research tools that contravene platform TOS – if platforms or users are not aware that data collection is taking place, then the likelihood of researchers facing legal action is reduced significantly.
- **The use of third-party tools or deceptive methods can disincentivise researchers from publishing their work and/or incentivise them to be vague about research methodologies.** This in turn has a negative impact on the quality, comparability and replicability of online research into key societal issues and can engender further distrust between platforms and the wider research community.

Key Findings: Extremist Communities on Telegram, Discord and Odysee

- On Telegram, ISD researchers identified six German-language private channels and 80 private groups associated with far-right extremism and harmful conspiracy theories, with 12,049 subscribers in the largest channel. Our findings suggest that even these larger, supposedly private spaces are used for harmful activities, such as calls for violence against politicians, spreading false information and coordinating offline mobilisation.
- On Discord, ISD researchers identified 31 English-language Catholic extremist servers with a total of 9,585 followers, and 16 Islamist extremist servers with a total of 4,757 followers. In these servers, our research found significant amounts of extremely hateful content that targeted LGBTQ-communities, Jewish people and women. Both Catholic and Islamist extremists expressed anti-democratic views, called for the establishment of totalitarian religious states and shared violent content, including support for terrorists and terrorist organisations.
- On Odysee, ISD researchers identified 8,690 French-language far-right royalist videos, 6,035 neo-fascist videos and 4,084 Catholic fundamentalist videos. The analysis of these videos points towards the presence of severe anti-democratic ideologies within French-speaking communities on Odysee. This includes material denying the Holocaust and venerating National Socialism, both of which are potentially criminal offences in France.

Recommendations

Based on the findings, implications for researchers and broader policy context outlined in this report, we have provided a series of overarching recommendations for all online platforms, policy-makers and regulators, and civil society and academic researchers to address the range of technological, ethical and legal, and fragmentation barriers that we encountered during our research.

We acknowledge that some of our recommendations will create additional work for platforms, especially smaller platforms with fewer resources or technical capabilities. However, at present, platforms provide effective functionalities for facilitating the communication and coordination of harmful actors while only offering restricted or poor functionalities for public interest research into these spaces. We argue this imbalance must be addressed.

Technological Barriers

- Platforms should provide permissioned data access (e.g. through APIs) to third-party researchers conducting public interest research. These should be accompanied by clear, consistent and accessible documentation that includes guidance on the types of data that can be collected and sufficient limits on the collection of sensitive personal data to protect user privacy.
- In future regulation, policy-makers and regulators should require data access that is accurate, reliable and sufficient for public interest research from all platforms. This would include the wide range of smaller- and medium-sized online platforms rather than just focusing on those that are currently the largest.
- Academic and civil society researchers should share effective data collection approaches or tools, as well as any lessons learnt, for accessing the growing range of platforms across the evolving online ecosystem.

Fragmentation Barriers

- Platforms should provide data access in a systematic way that enables researchers to reliably access accurate data from across public areas of the platform, rather than requiring researchers to manually identify online spaces or communities of interest. Tools for systematic search, whether platform-native or third-party, should demonstrate reliability, coverage and accuracy (e.g. through third-party review).

Ethical Barriers

- Platforms should determine a reasonable limit for the number of members that can participate in private groups and channels and declare online spaces with large audiences over a certain threshold as public. Content with no reasonable privacy expectations (e.g. content posted on public pages) should be made available via vetted API access.
- Policy-makers could consider introducing requirements for companies to clarify which areas of their platforms are truly public or private, and set reasonable thresholds for the number of users that can participate in private online spaces if platforms are unwilling or not incentivised to do so voluntarily.
- The research community should work to formalise ethical approaches to researching public, semi-private and private online spaces, in-line with the potential severity of the risks such spaces could pose, while also respecting users' right to privacy.

Legal Barriers

- Platforms should provide clear TOS, not only in regard to the types of content and activities they allow, but also to how they apply to researchers accessing data. These should then be enforced consistently.
- Policy-makers should introduce legal protections for researchers conducting public interest, privacy-respecting online research.
- Academic and civil society researchers should consider opportunities to share or pool expertise on the legal implications of platform data access.

Glossary

Alt-tech describes social media platforms used by groups and individuals who believe their political views have made major social media platforms inhospitable to them. This includes platforms built to advance specific political purposes; libertarian platforms that tolerate a wide range of political positions, including hateful and extremist ones; and platforms which were built for entirely different, non-political purposes like gaming.

An **API (Application Programming Interface)** is a software intermediary that allows two applications to communicate with each other. APIs have a huge range of uses, but in the context of this report, they allow researchers to access certain data from some online platforms via requests. As an intermediary, APIs also provide an additional layer of security by not allowing direct access to data, alongside logging, managing and controlling the volume and frequency of requests.

Conspiracy theories are attempts to explain a phenomenon by invoking a sinister plot orchestrated by powerful actors. Conspiracies are painted as secret or esoteric, with adherents to a theory seeing themselves as the initiated few who have access to hidden knowledge. Supporters of conspiracy theories usually see themselves as in direct opposition to the powers who are orchestrating the plot; these are typically governments or figures of authority.

Crowdsourcing and surveying methods involve users of online platforms voluntarily reporting particular forms of content to researchers through mechanisms such as browser plug-ins or reporting forms for users.

ISD defines **disinformation** as false or misleading content that is spread with the intent to deceive or secure economic or political gain, and which may cause public harm. When referring to such content that is spread unintentionally, we will be using the term **misinformation**.

Encryption refers to the process of encoding information so that it is rendered incomprehensible to everyone except specified receivers.

Ethnographic research methods involve deep and sustained involvement with a community. Instead of relying on data-collection technologies, researchers may take a more human approach by joining, participating in and observing online spaces as forms of community.
ISD defines **extremism** as the advocacy of a system of belief that claims the superiority and dominance of one identity-based 'in-group' over all 'out-groups'. It advances a dehumanising, 'othering' mindset that is incompatible with pluralism and universal human rights.

We define **fragmented platforms** as those where online content is theoretically accessible, without technological or ethical barriers, but nevertheless cannot be searched quickly or systematically, for example, via an API. Relevant content must therefore be found manually amid vast amounts of other material.

We use **harmful content and behaviours** to refer to a broad spectrum of online activities that can have a negative impact on human rights, society and/or democracy. These can include targeted harassment of individuals, incitement of violence against a particular group or the spreading of disinformation and harmful conspiracy theories. In some instances, the risk of harm may be intrinsic to the content itself, with the risks exacerbated by amplification; in others, the harm may be caused by aggregate patterns of behaviour rather than the nature of the content itself. Depending on the geographic and legal context, different forms of harmful content and behaviours may or may not be illegal. Depending on the platform, these also may or may not be covered by a company's 'Community Guidelines', standards or rules.

Hate is understood to relate to beliefs or practices that attack, malign, delegitimise or exclude an entire class of people based on protected characteristics, including their ethnicity, religion, gender, sexual orientation and disability. Hate actors are understood to be individuals, groups or communities which actively and overtly engage in the above activity, as well as those who implicitly attack classes of people through, for example, the use of conspiracy theories and disinformation. Hateful activity is understood to be antithetical to pluralism and the universal application of human rights.

Open Platforms are social media platforms on which content is visible to general users without further verification and often accessible via search engines. By contrast, content on **closed platforms** will not be easily accessible via search engines and often requires additional authentication or an invitation. Platforms will often contain open and closed elements; for example, Facebook has public (open) and private (closed) groups.

Systematic search methods use technology to extract large amounts of data and metadata directly from online platforms. Data might include, for example, the content of online text, connections between online accounts and metadata, such as the time or geographical location of posts. Many social media platforms also make data easier to access by providing APIs that allow researchers to directly access various forms of data from platforms without needing to build their own code from scratch. The development of AI-based approaches has also allowed for ever more sophisticated analysis methods; for example, natural language processing (NLP) is increasingly used to detect trends, sentiments and entities mentioned across vast quantities of online text.

Introduction

Harmful actors use an ever-expanding range of digital spaces to spread harmful ideologies and undermine human rights and democracy online. Understanding their evolving ideas, online networks and activities is critical to the development of a more comprehensive evidence base to inform effective and proportional efforts to counter them. But generating that evidence base can challenge the technical capabilities, resources, and even ethical and legal boundaries of research. We are concerned that these issues may be worsening just as the options for spreading harm online increase.

This difficulty in conducting digital research systematically, ethically and legally results in a situation where trade-offs have to be made between competing priorities, including the desire to understand and mitigate harmful content and behaviours online, the preservation of privacy and the adherence to legal agreements. We will argue in this report that this does not need to be the case; solutions are available, and actions should be taken as soon as possible to ensure a future-proof scenario in which researchers have the tools to monitor, track and analyse harmful content and behaviours systematically, ethically and legally.

The report outlines the findings from the research phase of a project by the Institute for Strategic Dialogue (ISD) and CASM Technology; it is funded by Omidyar Network. The aim of the project is to identify and test research methodologies for monitoring and analysing small, closed or hardly moderated platforms. The Phase I report from this project outlined the findings of an initial platform and methodology scoping exercise.[1] The focus was on identifying the key barriers posed by these platforms to researching and mitigating harmful content and behaviours, as well as reviewing existing research methodologies and tools to address these barriers. We developed three overarching categories of barriers from this exercise: technological, ethical and legal, and fragmentation.

- **Technological barriers** may be posed by, for example, encryption, AI-generated content, blockchain, decentralised platform structures or content formats (particularly audiovisual ones).
- **Ethical and legal barriers** may arise from, for example, expectations of privacy, legal restrictions, platform TOS that prohibit legitimate public interest research and difficulties in obtaining informed consent. In some cases, platforms may present ethical barriers without legal barriers and vice versa.
- **Fragmentation barriers** are present on platforms where online content is theoretically accessible without substantial technological or ethical barriers but cannot be searched quickly or systematically, for example, via an API. Relevant content must therefore be found manually amid vast amounts of other material.

In that report, we argued that the migration away from mainstream social media platforms towards online spaces offering less moderation may make it harder to research and address harmful activity online. This argument was based on the observation that many of these platforms present technological, ethical and legal, or fragmentation barriers for researchers, thereby limiting their ability to study harmful communities systematically.

This Phase II report builds on that analysis. It provides applied examples and evidence for the limitations and dilemmas encountered by researchers. In three small research case studies, we seek to apply different methodological approaches to analyse platforms that primarily present technological, ethical and legal, or fragmentation barriers. We examined online communities on Telegram, Discord and Odysee in German, English and French language respectively. The aim of this exercise was to expand the research field's understanding of which methodologies are applicable to these online spaces and provide applied examples of the types of barriers researchers might face when trying to access data from online platforms.

As will become clear throughout these case studies, these barriers often overlap or are interconnected, combining in different ways on different platforms. For example, legal considerations may make it impossible to overcome fragmentation barriers through systematic search methods; technical features may also create ethical barriers; or vague TOS or the lack of an API can cause uncertainty around what is technically feasible or legally permissible. We chose to focus on three platforms.

The lessons learnt from this research will feed into Phase III of the project, which will seek to inform practical, technical and regulatory solutions to data access and transparency for these types of online spaces without impinging on the rights of users. We will share and discuss our findings with relevant stakeholders, including research experts, policy-makers and regulators, and technology company representatives whose work touches on data access and transparency. Based on our findings, we will also consider how the legal and regulatory landscape may need to adapt to keep pace with the growing range and technological variety of online platforms, while also respecting and protecting vital rights to privacy, security and anonymity online.

Indeed, negotiating how to protect the right to privacy while conducting public interest research on harmful communities and behaviours has proved to be one of the key challenges of this project; ethical and legal concerns around legitimate expectations of privacy are flagged throughout this report. Unfortunately, we lack a common definition of private spaces, making it difficult for researchers to determine which spaces to treat as private versus public. At the same time, this absence of a definition allows platforms to limit transparency and access to private spaces that are easily accessible and arguably public. As we have previously argued, factors such as the size, the purpose, the accessibility and the nature of relationships between users of a channel or a community should be taken into account when making assessments about public or private spaces.[2] There is also a danger that this uncertainty could undermine the case for the encryption of genuinely private online spaces and means of communication.

The report first presents the three research case studies, starting with an introduction to each platform and its key functionalities, as well as how they have previously been exploited by harmful online communities. It then explains the specific ideologies and online communities we examined on each platform and outlines our rationale for selecting them. Subsequently, we document the findings of each case study, with a focus on the harms presented by these different communities. We also reflect upon the suitability and limits of the methodologies we identified during Phase I to overcome the anticipated barriers alongside any additional or unexpected barriers encountered during the research itself. This includes a discussion of methods that were technically feasible but not pursued due to ethical, security or legal reasons. The final section considers whether the three types barriers and research methods developed during Phase I stand and the implications this has for data collection, future research efforts and efforts to counter harmful content and behaviours on these platforms, with a particular emphasis on implications for digital policy and regulation.

Case Study 1: **Telegram**

Key Findings

- Telegram allows users to create private groups and channels. Content from these cannot be systematically searched, presenting fragmentation barriers. As these spaces are designated as private and joining them may require researchers to be deceitful about their identity, they also present ethical barriers.
- Nonetheless, these groups may have thousands of members, straining the definition of private online spaces. There is also evidence they are used for harmful purposes, including plotting potentially violent activity against German politicians.
- We therefore recommend Telegram declares groups and channels with audiences over a certain threshold as public and consequently, subject to the platform's TOS. We also recommend Telegram moderates these spaces more proactively and accepts takedown requests related to illegal content or activity within them.

For the case study looking at the German-speaking communities, ISD undertook research into far-right and conspiracy communities on Telegram. While data availability from groups and channels that are designated as public tends to be relatively comprehensive, there are ethical and fragmentation barriers around researching groups and channels designated as private by the platform; indeed, whether they are in fact private spaces can be disputed given their size and purpose. We therefore attempted to use a quantitative link analysis to identify high-risk groups and channels designated as private, followed by a qualitative ethnographic approach to analyse these communities.

Like many other social media platforms, Telegram does not define a specific numerical threshold for what is considered private. Instead, it empowers the admins of groups and channels to declare whether they are private regardless of their size. This has several key implications for how the platform is used. For example, as Telegram's TOS only apply to channels designated as public, this creates a loophole that allows users to get around the platform's already limited content moderation efforts. While Telegram is aware of this and states that in 'content disputes' private channels with publicly available links are treated as public channels[3], there is no evidence that the platform pro-actively moderates private channels. The functionality of private channels also enables admins to generate funds through paywalls that limit access to these private spaces.

There are also problematic consequences for researchers. Designating a group as private allows admins to limit the visibility of historical messages to the most recent hundred for new users joining a group or channel and/or restrict the saving or downloading of content. While it may be technically possible to access these messages via the Telegram API as a member of the group or channel, this could contravene the platforms' TOS and also increase potential ethical risks. Admins could also choose to disable this access to historical messages, and may be alerted to the presence of researchers using this approach, creating further challenges when researching online extremist communities. These barriers combine to limit researchers' ability to determine whether there is harmful content in private groups or channels in a systematic way (which we describe as a fragmentation barrier), even when their size and purpose would strongly suggest that users would be unlikely to have legitimate or realistic expectations of privacy. As such, we studied the content and behaviours in groups and channels using ethnographic methods, which do not rely on systematic large-scale data collection (as discussed in our Phase I report).

However, ethnographic studies of private groups on Telegram face ethical barriers due to the questions that often need to be answered to enter them (e.g. new members are asked to introduce themselves and occasionally provide some detailed personal information). These questions may require researchers to employ some level of deception (though often significantly less than on Discord, as discussed later). Finally, admins vet members depending on their behaviour and may only share links to more clandestine groups with members that they consider trustworthy. This means that researchers would be required to create a deceptive persona that behaves in a certain way to gain credibility and access to smaller harmful groups. The result is a difficult trade-off: the more concerning a private group appears from

the outside, the more legitimate it would appear to use deceptive means to enter it. On the other hand, it becomes harder and harder to justify increasingly deceptive measures, especially when attempting to enter spaces in which the line between private and public is blurred.

These ethical barriers and considerations prevented us from entering smaller groups and channels that would have required significant deception; our findings were limited to larger communities to which it was easier for researchers to gain access. Our findings suggest that, despite their relative accessibility, even these larger spaces are used for harmful activities, such as calls for violence against politicians, spreading false information and coordinating offline mobilisation.

The following sections outline the background and functionality of Telegram, the research approaches that we scoped and employed, the limitations encountered and the findings of our analysis.

Platform Overview

Telegram is a messaging app with some functions of a social media platform. According to its own statistics, it has 700 million active monthly users as of September 2022.[4] It was founded in 2013 by the Russian businessmen Pavel and Nikolai Durov. While the initially declared goal of the founders was to allow secure communication without government surveillance, the messenger has developed over time into a platform popular with a wide range of actors, including dissidents and activists in undemocratic or authoritarian settings. Telegram channels – chronological feeds of messages to which other users can subscribe – have become popular as a media platform for actors interested in spreading their content to a wide audience. Channels vary from more traditional media outlets to state propagandists and conspiracy influencers.

Telegram's TOS only prohibit the spread of illegal pornographic content and the promotion of violence by public channels and bots, as well as spam and fraud.[5] Although the platform has deleted terrorism-related channels, including in cooperation with law enforcement, it dedicates little to no moderation efforts to other illegal and harmful activities.[6] This has made Telegram particularly attractive to extremist and hateful actors banned from other social media platforms.[7]

Key Functionalities

Telegram users have an option to create and join public and private groups and channels. Public ones can be found via the Telegram search function, and anyone can join them (as well as see their content without doing so). Private groups and channels can be joined only through invite links created by admins or by admins adding new members manually. Although Telegram's defines a group or channel as private if it cannot be found by the search function and joined without an invite link, the actual level of privacy depends in practice on the actions and choices of their admins.[8] For example, admins can create different invite links with different limitations; for each specific link, they have the option to decide whether a person joining the group or channel needs additional approval from admins. They can also set a limit on how long the link is valid for and/or the number of users who can join the group or channel with it. Finally, the space in which a link to a private group or channel is posted can also affect the privacy and difficulty of joining it; links to private groups can be shared in public groups and channels with a high number of followers, or alternatively, admins can decide to share these links only in private messages or not to share them at all, instead adding new members manually by selecting them from their contacts.

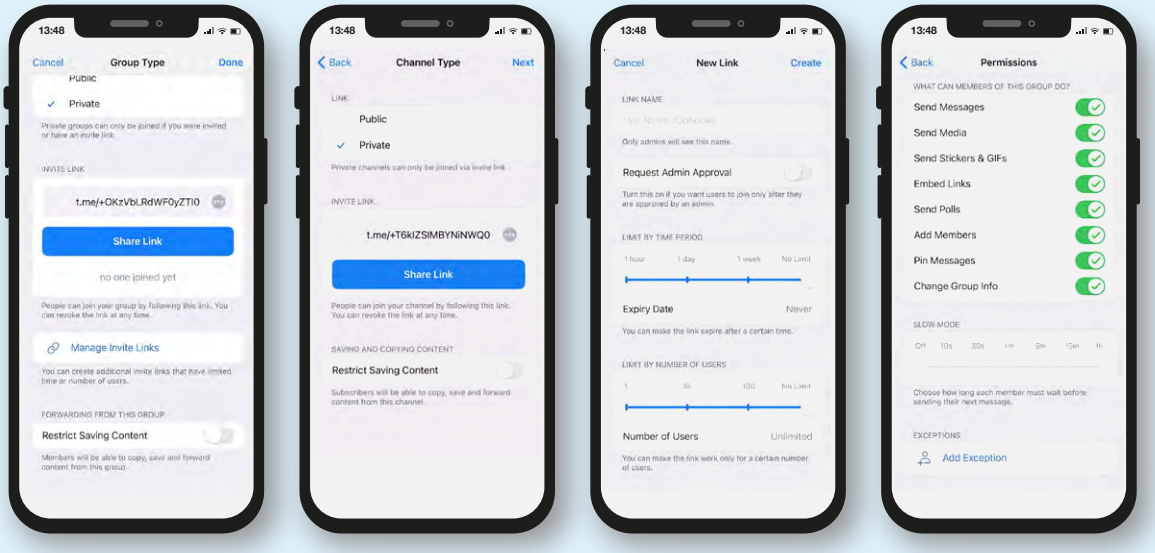


Figure 1: Telegram group and channel admin settings.

Private Telegram groups can have up to 200,000 members. Large groups that are private in line with Telegram's definition can hardly be considered truly private when they have several thousand participants who can all share invite links in public spaces. For example, TGStat, a popular Russian tool for analysing content on Telegram, contains both public and private groups in its database and regularly lists the most popular private channels for selected countries, showing both their names and invite links.[9] TGStat collects data from these private groups and channels, such as the number of participants or mentions of other groups and channels. This has already led to criticism, particularly from users in Belarus where authorities have detained people for merely subscribing to oppositional Telegram channels that are considered to be 'extremist' by the government.[10] TGStat claims that private groups and channels are in the database only if they are added by users or admins who are interested in statistics or if links to these spaces are shared frequently in public groups and channels.[11]

In addition to offering somewhat protected communication spaces, privacy settings on Telegram can also be used as a paywall. For example, German conspiracy influencer Oliver Janich[12] has several public Telegram channels as well as a private one, which is accessible only to paying subscribers.[13] For a slightly higher subscription fee, users can gain access to a private Telegram chat where they are offered the chance to interact directly with Janich. As of October 2022, it was possible to pay the subscription fee with cryptocurrencies. Janich was arrested in the Philippines in August 2022, but his various Telegram channels continue to function (maintained presumably by his team).

This means that many groups considered private by Telegram's definition in fact fall into a grey area similar to some other online communication spaces.[14] The degree of privacy on Telegram depends not just on technical functionality but also on the group's size and the decisions of its admins. The table below shows the spectrum of privacy in various types of communication spaces on Telegram.

⬅ More Public				More Private ➡
Public groups and channels.	Large private groups, links to which are posted in public channels and on other social media platforms.	Smaller private groups, links to which are posted only in other private groups for a limited amount of time. Admin approval might be required to join.	Private groups, links to which are never posted on any social media platform but rather are shared in one-to-one private messages. Admins might decide to choose members only from their trusted contacts.	Private messages between two users.

Figure 2: Spectrum of privacy for Telegram groups and channels.

Harmful Activity on Telegram

With little to no moderation efforts, Telegram has developed into a central social media platform for international extremists across the ideological spectrum, including German-speaking far-right, conspiracy and anti-lockdown actors and communities.[15] For example, in April 2022, German authorities announced that they had foiled a plot by extremist actors, connected via a Telegram group, to kidnap public figures, including the Health Minister.[16] In another case, the authorities opened a criminal investigation after journalistic research revealed that members of a local far-right private Telegram group in Dresden were discussing the assassination of Saxony's Minister President (the head of the Länder government).[17] Other journalistic research has documented numerous calls for violence in Telegram groups and channels.[18] Federal and regional criminal investigators in Germany and Austria have also discovered Telegram groups dedicated to selling drugs, weapons, falsified documents and stolen data.[19] While there is clear evidence of Telegram being used for illegal and harmful activities, systematic research about the scope and nature of these activities in private groups is still lacking.

Researching Telegram

Technological Barriers

From a technological point of view, data in private groups on Telegram is easily accessible once researchers have joined them. Members of closed groups can download conversation histories and save lists of group members via the 'export chat history' option in the group settings. This also provides access to older text and voice messages, as well as any pictures, videos and documents shared in the group. However, as outlined above, designating a group as private allows admins to limit the visibility of historical messages to the hundred most recent messages for new users joining a group or channel. Despite attempts to position itself as a platform focused on the safety and privacy of its users, Telegram does not offer end-to-end encrypted group chats. Encryption is available only for 'secret chats' between two users, and it is not a default option but has to be selected manually.
Some Telegram features, however, do present technological challenges for systematic research. Voice messages and other forms of audiovisual content, which are popular on Telegram, are particularly difficult to analyse in an automated way. In addition to this, Telegram offers group voice and video calls, which are impossible to analyse without joining them.

Another platform functionality available on Telegram (and many other platforms) that presents a challenge for researchers is so-called ephemeral messaging, where messages are only visible for a limited time. In one of the observed groups, admins temporarily activated the auto-deletion function for messages, making it impossible to collect the deleted messages with the export chat history function as well as limiting the time during which chat members could access them. While this type of functionality can play an important role in protecting users' privacy, it also presents a significant challenge to conducting online research.

Fragmentation Barriers

Another barrier to research on Telegram is finding groups of interest due to what we have called fragmentation. Private Telegram groups and channels are hidden from the platform's search function and from search engines. It is therefore impossible to find them using specific keywords. One approach is for researchers to conduct a systematic link analysis

in the public Telegram groups and channels they have already joined via automated collection, and then filter for links to private groups or channels, or search for links in databases or on other platforms. Alternatively, researchers could manually move from one group or channel to another and manually check whether links to other private groups or channels are posted there. However, this is much more time consuming than using the search function. Furthermore, links to some clandestine groups might never be posted in public spaces and would thus remain unavailable.

Ethical Barriers

Studying harmful activities online, particularly those by violent extremist communities, poses a set of ethical dilemmas discussed in detail by researchers and academics, as well as by lawyers, governments and regulators.[20] While there is a public interest in understanding the dynamics of radicalisation and analysing the communication of communities that plot violent attacks or are involved in other illegal activities, it can put researchers in a difficult ethical position when they have to balance this public interest with the rights of research subjects.

Such ethical considerations are a major complication when it comes to research of harmful content and behaviours in private Telegram groups where users expect that they are communicating without intrusion from external observers. If researchers want to obtain informed consent in order to fully respect the research subjects and their right to privacy, they need to be transparent about their identity and goals. They must ask all group participants to consent to the presence of a researcher in the chat and to the potential storage and analysis of their data and messages. This consent is likely difficult to obtain, especially in the case of extremist online chats. Knowing that they are being observed, members of these communities might also change their behaviour and avoid potentially illegal content or simply leave the group altogether. For example, users in conspiracy groups who may see academic institutions and think-tanks as a part of a cabal of global elites would be unlikely to welcome a potentially critical observer. Finally, risks to the researchers themselves must be considered; if their identity and institutional affiliation are revealed to potentially harmful actors, they might become targets for online abuse, harassment, doxing or even offline attacks.

If researchers find that conducting their work in a fully transparent way is either impossible or too risky, they might decide that the value of studying a particular private group outweighs potential concerns over privacy rights, unconsented research or some forms of deception. Researchers should also consider the costs of not studying potentially violent and harmful groups that could violate others' rights or prepare violent acts.

While some closed Telegram groups are easily accessible and de facto semi-public, others only have a handful of participants, and entry requires approval from the group's admins via asking questions of potential members. In the former case, researchers can likely remain as passive observers without using any active deception. In the latter case, they are likely to be asked about their motives for joining the group and must decide if they want to provide deceptive answers to gain access.

Research Methodology

In order to find invite links to private groups and channels, ISD researchers joined a seed list of public channels. This list of 253 German-language far-right and conspiracy public channels was compiled during work on another ISD project on extremism in Germany and updated for this project.[21] It is worth noting that this approach is only possible for researchers that have access to existing and up-to-date lists of groups and channels, which can be challenging to create and maintain, representing a further barrier to research.

Messages published by these channels from 1 January 2022 to 18 August 2022 were collected, and links were then automatically extracted from the messages. ISD researchers manually checked links to Telegram's t.me domain (a common format of links directing to Telegram channels, groups and messages) in order to identify potential invite links. Without joining the groups or channels, it was possible to identify whether the link was active or not; to see the name of the private group or channel, its profile picture and description (if available); and the number of members (groups) or subscribers (channels).

The result of this analysis was a list of 80 private groups and 6 private channels. The number of members or subscribers varied between 8 members in the smallest group and 12,049 subscribers in the largest channel.

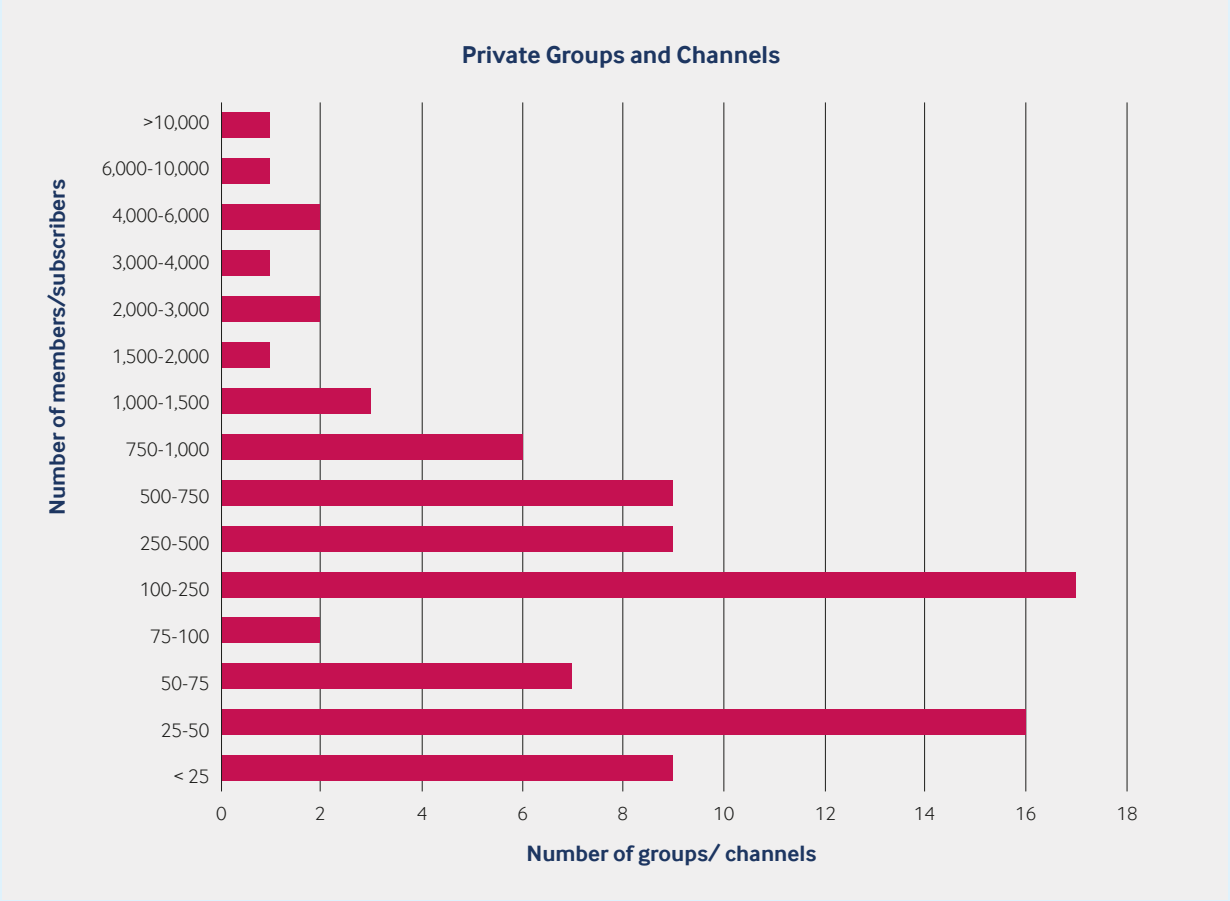


Figure 3: Size of private groups and channels in the analysed sample.

While some invite links in the initial dataset had been shared frequently, others had been shared only once and received a small number of views. The most widely shared invite link to a private group connected to a conspiracy Telegram channel which serves as a discussion space for the channel's subscribers; it was shared 1,843 times and received over 35 million views. This group had 2,852 members, which can hardly be considered a fully private communication space.

Thematically, as far as it could be judged from names and some of the available descriptions, the private groups and channels in the dataset were dedicated to the following topics:

1. Mobilisation of protests in specific German regions and towns. While several groups listed the name of the protest in the group description or names, others used general phrases such as 'We rise up' or referred to 'Monday walks' – initially, a form of democratic protest in former East Germany, the name of which has been recently appropriated by far-right actors and Covid-19 sceptics.
2. Discussions on the Covid-19 pandemic. From the descriptions, these groups united anti-vaxxers and Covid-19 sceptics, although some names and descriptions were more euphemistic, mentioning 'healthy free thinkers' or protecting children (presumably from vaccinations and other Covid-19 measures).
3. Regional help for truckers from Russia. These groups were set up after the start of the Russian invasion of Ukraine. Invite links to them were widely shared by public far-right and conspiracy channels.
4. Specific events other than protests, such as get-togethers, or commercial esoteric events, such as seminars on spiritual growth.
5. Conspiracy theories, including QAnon.
6. Esoteric topics, such as alternative medicine or spiritual growth.
7. Regional networking without any reference to the topic of discussion.
8. Other topics or groups with unclear purposes.

When deciding on which to join, we prioritised either the most popular groups and channels with a large number of members or subscribers, or those that were likely to be most extreme from their handles and descriptions (including known extremist groups or those that were likely to share harmful conspiracies). However, it was not always possible to understand the topic and potential level of harm correctly without joining the groups or channels first. For example, our dataset contained links to a private group and a private channel that were shared by far-right channels and looked like they were related to a neo-Nazi band. In fact, these were related to a hoax by anti-fascist activists, which was designed to show how far-right music is being spread on music streaming platforms.[22] Another private channel in the dataset had a deceptive name, 'Animal protection is a matter of honour'; as the link to the channel was shared by prominent German far-right and conspiracy influencer Attila Hildmann, we decided to join it. The channel proved to be one of Hildmann's reserve channels created to circumvent restrictions imposed by Telegram on his main channels.[23]

Entry Requirements

In several larger groups (several hundreds of members), researchers had to solve simple CAPTCHAs (a test to determine whether a user is human) on joining to be able to post messages.

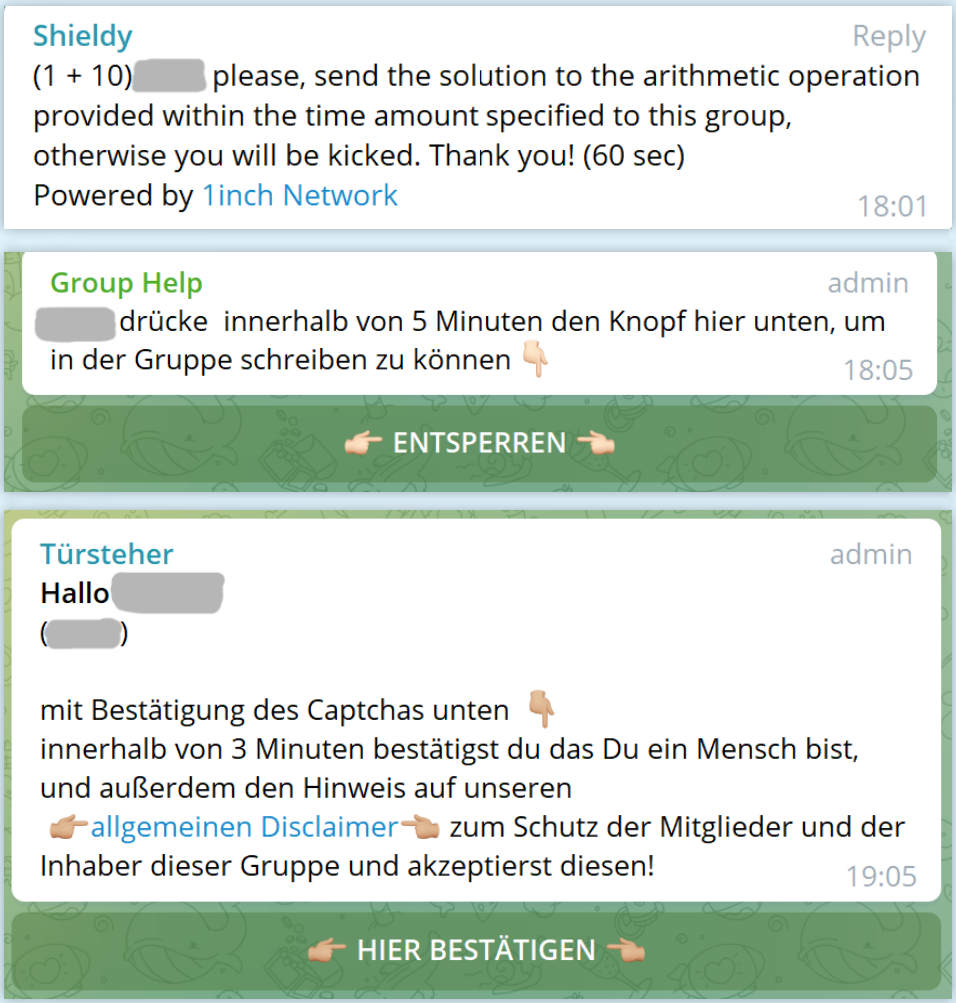


Figure 4: CAPTCHAs in private Telegram groups.

In one group with over 2,000 members, a bot asked new members to answer within three minutes why they were joining the group with three possible answers: 1) 'No idea', 2) 'Is this important?' and 3) 'I am curious'. Those that chose the first or second options were automatically kicked out of the group and prevented from joining it again.

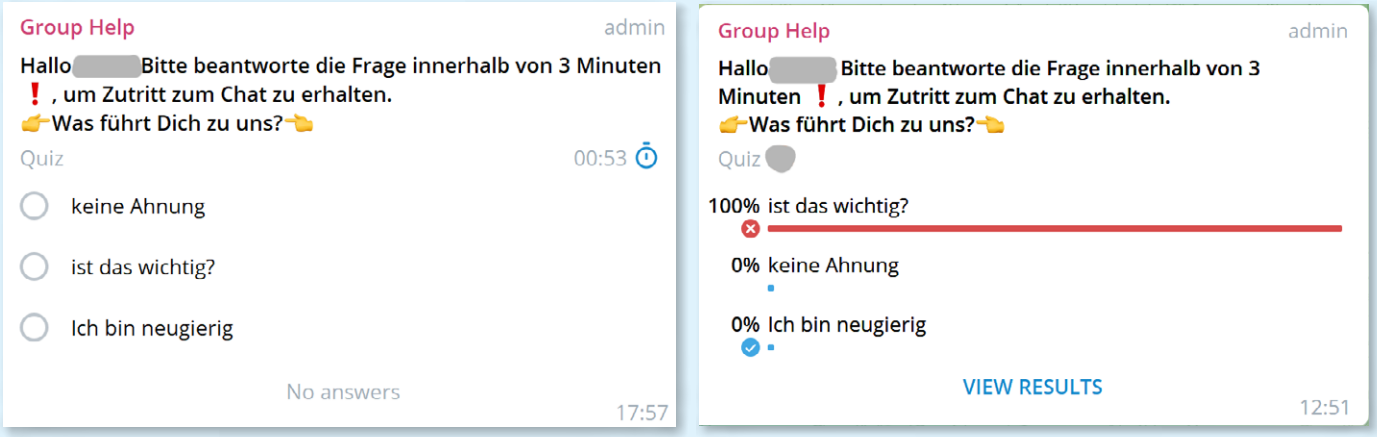


Figure 5: Questions asked by a bot as an entry barrier in a Telegram group.

In one of the smaller closed groups – a regional chat with a declared goal of networking former soldiers – new members were asked by a bot to introduce themselves and give some personal information. These regional groups were connected to a public Telegram channel for veterans, and which had more than 14,000 subscribers. It had been regularly publishing pro-Kremlin disinformation about the war in Ukraine, including from Russian state outlet RT, as well as false and misleading information about the Covid-19 pandemic and calls for protests against the German governments' public health measures and sanctions against Russia. The channel was connected to a public group with more than 600 members, which also served as a comment function of the channel.[i] In this public group, members were posting calls for violence against politicians, calls for buying arms and for an 'uprising' against the German government instead of 'pointless demonstrations'. Notably, Telegram's TOS do not explicitly prohibit the promotion of violence in any groups, regardless of whether they are public or private. According to Telegram, 'all Telegram chats and group chats are private amongst their participants. We do not process any requests related to them.'[24]

Figure 6: Calls for violence and 'uprising' in a public group for veterans.

i In order to enable comments on posts in a Telegram channel, channel admins have to connect it to a group.

After joining the public group for veterans, researchers were able to see invite links to regional closed groups for networking, which were significantly smaller (members ranging from 7 to 100 as of September 2022). The decision to join small closed groups such as these presents an ethical dilemma; however, in this case, researchers decided that these communities might be potentially harmful and violent, given the calls for violence in the public chat and the fact that the groups and channels were targeting former military personnel. The researchers joined one of the regional groups in which new members were asked by a bot to introduce themselves, name the first digits of their postcode and give information about their service time in the military. According to the rules of the group, new members who do not introduce themselves within 48 hours are kicked out of the chat. These rules, however, were not applied strictly in practice; researchers were able to stay in the group even though they decided not to give any deceptive answers or to write anything in the chat.

Overall, the barriers for joining large private groups, links to which were shared publicly, were relatively low. Researchers were often able to join groups and stay in them as passive observers without any active deception. For this research, we decided not to join any smaller groups without strong evidence of their involvement in potentially violent activities. We also refrained from the kinds of active deception which might have been required for the researchers' accounts and personas to be accepted as insiders in communities and therefore be invited to the more clandestine groups. This meant that we could only assess relatively open groups.

Analysis of Telegram Communities: Key Findings

We had to decide where to focus our ethnographic work among the private groups and channels collected by our link analysis. Given the evidence we found of calls to violence and illegal activity, particularly by German-language far-right, anti-lockdown and conspiracy groups, we focused our research on known extremist groups and those that spread calls for violence, offline protests or might potentially plan violent activity. Overall, researchers joined 28 private groups and 2 private channels.

During the observation period (August–October 2022), the groups were repeatedly posting calls for protests or a 'general strike' against any remaining Covid-19 related public health measures, Covid-19 vaccinations and increasingly, the rising cost of living and sanctions against Russia. Users in the groups frequently forwarded messages from conspiracy or far-right channels, posted links to alternative media or Russian state media, such as RT or RT DE. After the EU imposed sanctions on Russian state media following the full-scale invasion of Ukraine, users in the observed groups shared tips on how to circumvent the sanctions and posted links to alternatives RT domains. Members of the groups were also posting flyers about upcoming events and pictures and videos from events taking place in their regions, as well as in other regions, presumably to reinforce the feeling of urgency and create an impression of wide-spread protests by like-minded people.

To mobilise for protests, authors of posts frequently appropriated the language of anti-authoritarian resistance, stating that they were protesting for 'freedom', 'human rights' or 'peace'. Contrastingly, the German state was portrayed as a dictatorship or 'Corona regime'. Members of the groups were spreading false information that the government was planning to use the army against protesters or to shoot at them. The German government was also portrayed as directed by external forces or a 'global finance mafia' (an antisemitic dog whistle).

Figure 7: Flyers and photos of protests shared in private groups.

For one of the private groups, ISD downloaded the whole chat history to test possibilities for automated analysis. This was the group of a regional extremist group under observation because of suspected 'anti-constitutional delegitimisation of the state' by the regional Office for the Protection of the Constitution (a state-level domestic intelligence service) in Saxony-Anhalt.[25] This new category was introduced after anti-lockdown protests in order to deal with actors who delegitimise and sabotage state and public institutions.[26] After joining the group, researchers were able to manually save the chat history from a selected date (in this case, 1 January 2022). Between 1 January 2022 and 3 October 2022, 767 unique users posted 22,165 messages in the group. The volume of messages was particularly high from January to March, with most messages at that time focusing on opposition to public health measures or the beginning of the full-scale Russian invasion of Ukraine. The group became less active in May and June and then the level of activity increased again in the autumn (though not to the same extent as earlier in the year).

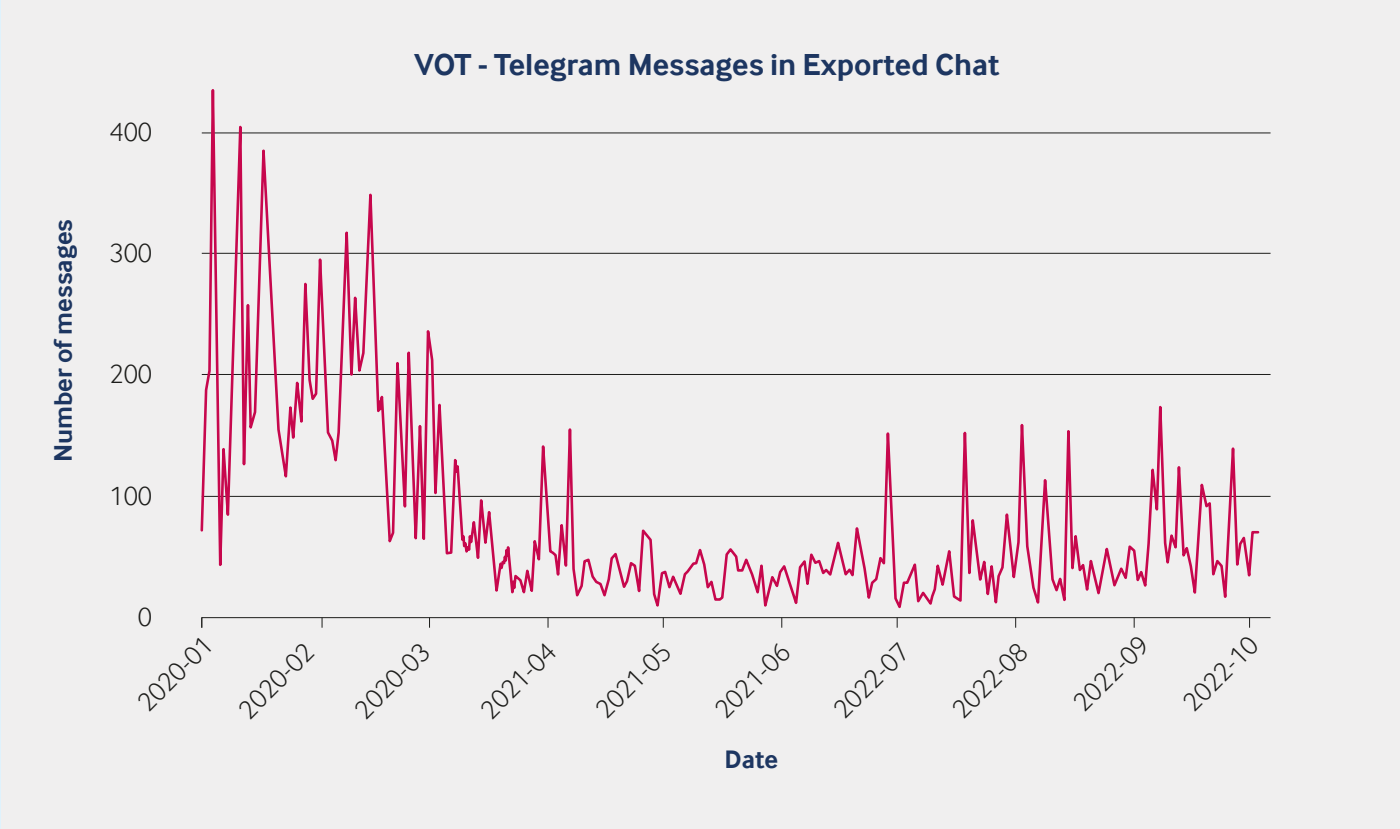


Figure 8: Volume of messages in the exported group's chat.

The majority of messages in the chat were posted by a small pool of highly active users. The 76 most active users (10% of all unique users), defined as those who posted 90 or more messages, were responsible for 78.7% of all message output in the group. The three most active users posted over one thousand messages each. Contrastingly, over half of all unique users posted just three or fewer messages during the observed period.

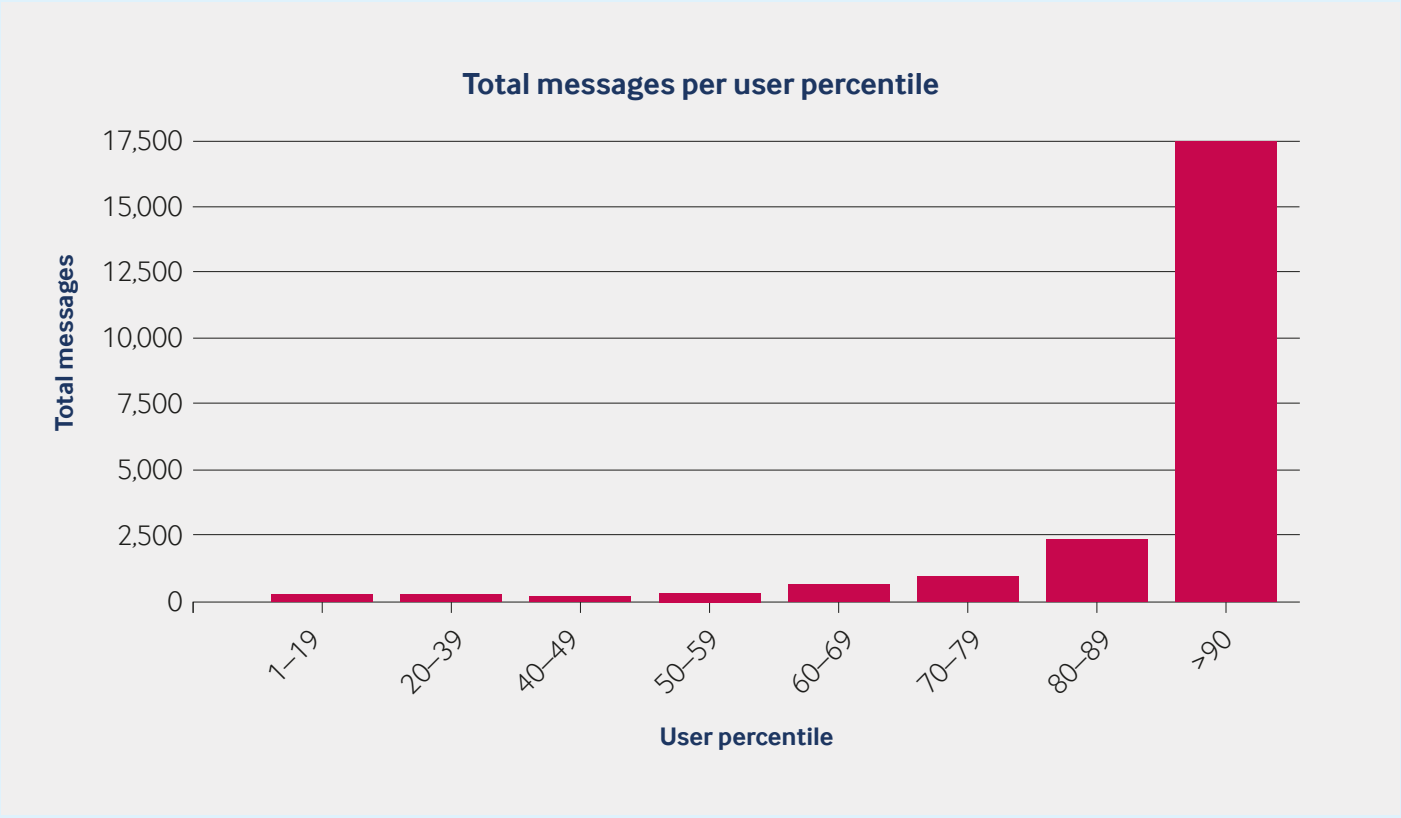


Figure 9: Total messages per user percentile.

Members of the group also posted a significant amount of audiovisual content that was difficult to analyse in an automated way. Overall, the chat history, starting from the oldest message, posted on 28 April 2020, through to 20 October 2022, contained 6,428 pictures, 3,128 videos, 179 voice messages, 85 audio files and 603 other files like PDFs and .doc files. The volume of audiovisual data made a thorough analysis of the files for just this one group a significant challenge.

Findings and Recommendations

Telegram remains a platform without a comprehensive system of moderation. Although the platform sporadically removes illegal content or content promoting violence, these actions remain inconsistent and insufficient. Private groups provide an additional loophole for circumventing legal requirements to remove illegal content that were put in place in countries like Germany. Content posted in private Telegram groups cannot be detected in an automated way via external search engines or Telegram's own search function or API. Instead, groups and channels must be found through systematic analysis of invitation links, which can be laborious (and therefore expensive from a resourcing perspective). It may be particularly challenging for researchers without existing lists of harmful groups or knowledge of where to find them.

Researchers must join private groups to get access to their content, posing an ethical dilemma. In order to access to the most clandestine and harmful groups, researchers would have to use active deception (i.e. behaving like other members of the groups and posting hateful or extreme messages) or pay subscription fees or donations to harmful actors to be accepted as a trusted member of the community. While these deceptive methods are sometimes used by journalists[27] or state law enforcement and intelligence services, we decided to refrain from paying extremists or behaving in a deceptive way for ethical reasons.[28]

Due to fragmentation and ethical barriers, our research of harmful communities was therefore restricted to larger private groups, links to which were shared in public channels. We found that some of these larger communities strain the definition of private due to their size and the fact that invite links were actively shared in public channels. Even these semi-public digital spaces were used for calls for violence, the spreading of harmful conspiracy theories and offline mobilisation.
The fact that Telegram's rules do not apply to either public or private groups (that can have up to 200,000 members) or to private channels makes these spaces an attractive online venue for harmful actors to spread calls for violence and false information, as well as mobilise communities for offline actions.

To address these harms, Telegram should determine a reasonable limit for the number of members that can participate in private groups and channels. Online spaces with large audiences over a certain threshold should be declared as public and therefore subject to Telegram's TOS. Telegram should also commit to enforcing these TOS more consistently and effectively in order to limit the volume of illegal or harmful content, behaviours and communities on the platform. Finally, Telegram should ensure that researchers can access and obtain data on these public online spaces, while also ensuring appropriate privacy and data protection safeguards for their users.

Case Study 2: **Discord**

Key Findings

- Discord presents similar challenges to Telegram – researchers must join a community (or server) to access content, which presents fragmentation barriers to systematic research as well as potential ethical and legal issues around deception and use of data.
- Similar problems in distinguishing between public and private spaces apply on Discord. In addition, the Discord's tagging and searching functionalities by which users find servers of interest seem to have various issues, producing inconsistent results when compared with third-party alternatives. Servers often mix harmful and more generally social behaviours, making locating relevant material even harder.
- Nonetheless, we were able to conduct a (limited) study of two different religious extremist communities: Catholic and Islamist. Across these two separate groups, we saw similar harmful themes, including anti-LGBTQ hate, antisemitism and calls for fundamentalist religious states.

For the case study looking at English-speaking online communities, ISD analysed those on Discord promoting two types of religious extremism: Catholic and Islamist. Discord primarily presents fragmentation barriers. To circumvent these barriers, researchers scoped two different methodological approaches: 1) systematically accessing data limited to servers of interest and 2) ethnographic research.

Discord poses fragmentation barriers in two distinct ways. Firstly, while searching and downloading messages via Discord's API is technically possible, this is only on a server-by-server basis, meaning that researchers need to know where to look for harmful content as well as be a member of the servers in question. Secondly, knowing where to look for relevant content or communities is made more difficult by the limited and non-transparent search functions provided by the platform as well as by third-party software.

While fragmentation barriers are the key obstacles on Discord, the platform also presents distinct legal and ethical barriers as well. As Discord's TOS ban the collection of user data via its API, gathering data in this way from servers would appear to break contract law and risk possible legal action by the platform, even in cases where there are no ethical concerns around gathering such data; these legal considerations therefore make it difficult to use systematic search methods to overcome fragmentation barriers. Correspondingly, ethical barriers limit the extent to which researchers can employ qualitative methods to legitimately study harmful communities on Discord despite an absence of legal barriers to doing so. The questionnaires that usually need to be answered to enter groups may require significant deception or the promotion of problematic beliefs; this is especially the case for smaller servers as it is not clear where the threshold for reasonable expectations around privacy begins.

These barriers combine to limit the scope of our research findings, which mainly rely on qualitative and non-systematic insights from servers where ethical considerations did not outright prevent us from joining in the first place. The findings nevertheless point towards a significant presence of hateful, anti-democratic and violent content within these communities.

The following sections outline the background and functionality of Discord, the research approaches that we scoped and used, the limitations encountered and the findings of our analysis.

Platform Overview

Discord was launched in 2015 as a free video-gaming platform and was designed to assist gamers in communicating with each other while playing. Since then, Discord's userbase has grown dramatically; it is currently estimated at 6.7 million active servers and 140 million monthly active users worldwide.[29]

Key Functionalities

Discord allows users to talk to each other in real time via text, or voice and video chats between users. Chat rooms – known as servers – can be created by any user for purposes that extend beyond gaming, including networking, organising events and competitions, thematic discussion, raids (organised campaigns to spam or troll other servers or users on other platforms), and collating and sharing content that is of interest to servers' members.

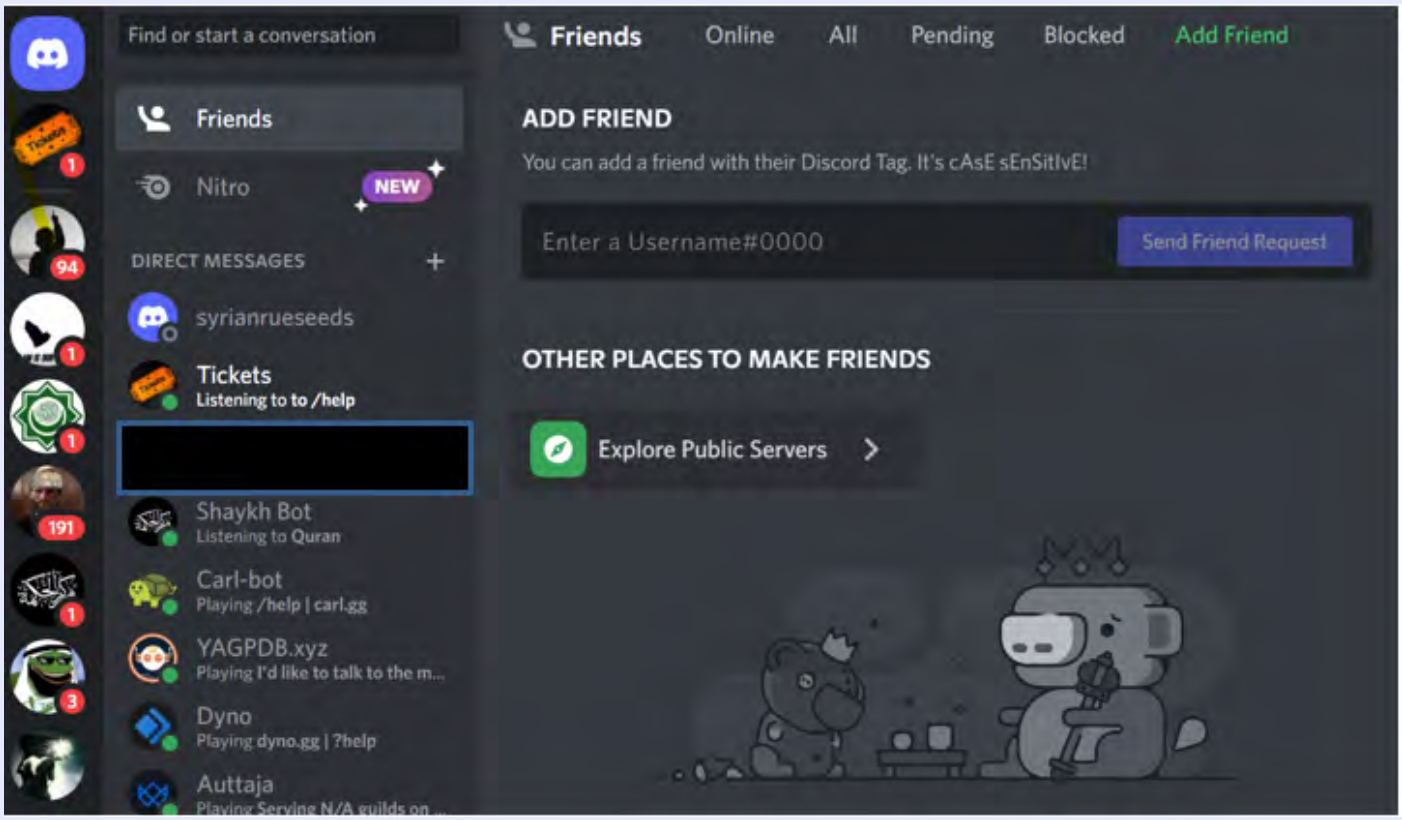


Figure 10: Discord user interface.

Many of these servers are private, but many are also public (though still require a username to join). The largest public servers can have hundreds of thousands of members.[30] While many public servers are dedicated to gaming or anime, some are dedicated to social or political discussion (with some explicitly discussing and drawing on themes and aesthetics that are similar to communities found on other platforms like 4chan).[31]

Harmful Activity on Discord

Despite the fact that Discord is primarily aimed at gamers and designed for non-political purposes, researchers have documented that the app has been used by various extremist groups as well; for example, in the run-up to the deadly white supremacist rally in Charlottesville, Virginia in August 2017 (during which an far-right activist murdered a counter-protestor when deliberately driving a car into a crowd), organisers used Discord to plan and coordinate the protest, as well as share ideological propaganda materials.[32]

Following reports about the far-right's use of Discord, and the events in Charlottesville in particular, Discord began to take a stricter approach vis-à-vis the presence of these movements on its platform.[33] In 2021, Discord claimed

to have removed more than 2,000 extremist servers.[34] Nevertheless, ISD research later in the same year found that Discord still acts as a hub for far-right socialising and community building even though there is little evidence that gaming itself played a role in serious strategies to radicalise and recruit new individuals on the platform.[35] Additionally, the research identified expressions of support for Atomwaffen Division and Sonnenkrieg Division (both designated terrorist organisations in the UK, Canada and Australia) in far-right Discord servers.[36]

While Discord appears particularly popular among adherents of the far-right, Discord is also used by members of a younger community of Gen-Z Islamist extremists who merge Salafi ideasii with alt-right memes and gaming subcultures.[37] Discord's role in a diverse range of online-subcultures adjacent to extremism and violence was further highlighted in 2022 by the Buffalo attack and the Highland Park Parade shooting; in both cases, the attackers' digital footprint contained messages left on Discord, including content that documented their attack planning and preparation.[38]

Researching Discord

Research Methodology

Discord presents both legal and ethical (primarily for quantitative and qualitative methods respectively when researching private groups) and fragmentation barriers (for both public and private groups) as research on the platform can only be done server by server and not in a systematic way. To circumvent these barriers, researchers scoped two different methodological approaches: 1) systematically accessing data limited to servers of interest and 2) ethnographic research.

As noted above, previous research by ISD on Discord has shown that both Islamist extremists and the far-right have a presence on the platform. ISD researchers considered if a comprehensive analysis of far-right groups could become the aim for this project as previous research on Discord hinted at a significant far-right presence on the platform; however, it was decided that a narrower and more focused study of specific sub-communities within the far-right could contribute more to the ongoing efforts by researchers to understand the diversity of extremist communities and radicalisation dynamics on Discord. Given reports of a growing interest within the US far-right in concepts like Catholic integralism and traditionalism, including among adherents of the far-right America First and 'Groyper' movements that have mobilised on Discord,[39] ISD chose to focus on hateful and violent content within Catholic integralist and traditionalist servers.iii This was supplemented with an analysis of Islamist extremist communities on Discord, allowing us to test which methodologies were applicable and which barriers researchers might face when trying to access data from harmful communities whose presence on Discord had been previously documented.

To identify Discord servers relevant to Catholic and Islamist extremism, we followed two steps. Firstly, researchers created keyword lists of vocabulary associated with Catholic and Islamist extremism online by using Discord's platform search functions, as well as Disboard (an open source tool that is not affiliated with Discord but allows users to find servers).[40] Keyword lists were created based on previous ISD research into Islamist extremism and a review of the existing publications and reporting around Catholic extremism in English speaking countries.

ii Salafism is a form of Sunni Islam whose followers advocate a return to the practices of the first three generations of Muslims (the salaf or 'ancestors') who lived immediately after the prophet Mohammed. Within Salafism, there are multiple currents which differ significantly in their interpretations of the holy scriptures of Islam and their implications for political action. Salafis are often subdivided into quietist Salafists, who reject political activism; political Salafis, who are actively engaged in transforming society according to their ideological ideas; and Salafi-jihadists, who use violence to implement a Salafi interpretation of Islamic law.

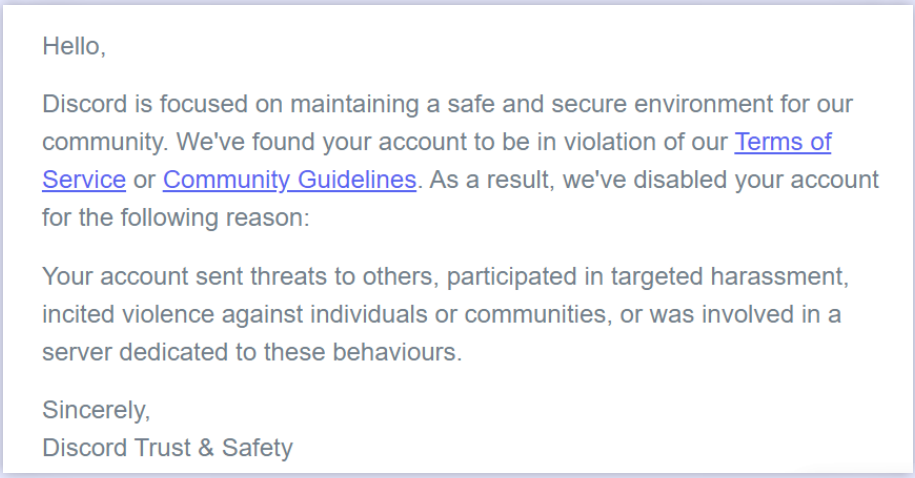
iii Catholic integralism is a trend within Catholicism that rejects the separation between secular and religious power. Since eternal ends are supposedly more important than temporal ones, it is argued by integralists that the teachings of the church should determine politics to serve the common good. Catholic integralism is not a recent trend, but over the past few years, these ideas have gained traction in English-speaking far-right (online) communities. There is some overlap with Catholic traditionalists who reject the Catholic church's partial embrace of modernity (especially around issues such as freedom of religion, gender roles and human rights, as well as around liturgy) following the Second Vatican Council between 1962 and 1965, with some traditionalists no longer recognising the legitimacy of subsequent popes and/or their teachings. For example, sedevacantists believe that the position of Pope has not been filled since the Second Vatican Council, while sedeprivationism holds that the Pope is legitimately elected but lacks authority due to doctrinal errors. Due to their anti-secular, anti-liberal and anti-pluralist stances as well as their rejection of the Second Vatican Council reforms, there can be some overlap between the distinct trends of sedevacantism and integralism. Catholic integralist ideas are discussed on a wide range of platforms, including on Discord.

Building on the first step, researchers then undertook a manual 'snowball' search on each platform (i.e. using initial cases to discover and expand into new cases). Researchers drew on servers that were recommended within extremist discussions, and 'partnership' servers in particular. Partnership chats within servers are used to list like-minded servers which promote each other to their members as they are believed to potentially be of interest. Partnership descriptions are generally much more extensive than the descriptions within Disboard and usually allow researchers to draw indicative conclusions about their likely ideological background.

Following these steps, researchers identified 31 Catholic and 16 Islamist extremist servers. To classify the ideology of servers, ISD researchers assessed their name, tags on Disboard and/or descriptions. In many cases, there is no one coherent ideology within a server but a mix of different voices that have significant disagreements while also discussing other more benign topics. It is therefore more accurate to describe the majority of the servers we investigated as communities in which extremist voices are prominent rather than extremist servers per se.

For this case study, two accounts were created. This was done to avoid suspicion from other users of relevant servers as it is possible to see in how many (and which) servers a user has common membership with others; for example, it would likely cause suspicion if researchers who are joining Islamist extremist servers are already members of Catholic extremist servers.

It is worth noting that our account created to research Catholic extremist communities on Discord was suspended two months after joining the identified servers; Discord stated that the account had been in 'violation of our Terms of Service or Community Guidelines.' While the specific reason was not given, it appears from the list of behaviours provided that the suspension was due to other members of these servers sending threats, participating in targeted harassment or inciting violence. Many of the servers analysed here would have met one or more of these descriptions, and no information was provided as to which server led to this decision or if the server itself was deleted. Several of the 31 Catholic extremist servers that ISD researchers had joined were no longer available via Discord's search function or Disboard when we concluded our data collection in August 2022 (though we are not entirely confident this necessarily means they have been removed by the platform).



Hello,

Discord is focused on maintaining a safe and secure environment for our community. We've found your account to be in violation of our Terms of Service or Community Guidelines. As a result, we've disabled your account for the following reason:

Your account sent threats to others, participated in targeted harassment, incited violence against individuals or communities, or was involved in a server dedicated to these behaviours.

Sincerely,
Discord Trust & Safety

Figure 11: Notification from Discord on removal of ISD research account.

As Discord has a publicly available API, we scoped the possibility of systematically accessing data for further analysis from the servers we identified as relevant through these steps. The following sections outline the fragmentation, legal and ethical barriers that we encountered.

Fragmentation Barriers

On Reddit, Facebook and other platforms, discussions in public groups can be accessed through the API. This means that a researcher can find mentions of relevant keywords quickly from across a range of public groups. A similarly wide-ranging functionality is not available on Discord. While searching and downloading messages via Discord's API is possible, this is only on a server-by-server basis. Some users have automated this to work at scale;[41] however, it appears that researchers would need to know in advance which channels they wish to search within. Given the huge range of channels on Discord and the fact that channels that host harmful content or behaviours are sometimes deleted and/or renamed, this can make systematic searching very challenging. To clarify, the issue here is not that the information is hidden – it would be easy to find if the researcher already knew where to look. Discord therefore presents fragmentation barriers to researchers.

Furthermore, it is not transparent how the search functions of Discord or Disboard work, how their outputs are determined, how comprehensive the results are and what results are missing. Disboard also groups servers together by tags that provide an indication of which servers could be relevant of harmful ideologies (e.g. ISIS, jihad, Taliban, Caliphate and Salafi). Each server name and description were manually reviewed to assess whether they were likely to be relevant to this project.

Two factors indicate that neither Discord nor Disboard's search functions are comprehensive. Firstly, they produce very different results to each other, with Disboard's generally being much more extensive. Secondly, even within Disboard's results, there are often unexplained discrepancies between the number of supposedly identified and actually displayed servers. It is therefore conceivable that their functionality could limit the accuracy of the data they provide, thereby potentially skewing our findings.

Ethical and Legal Barriers

Discord presents distinct ethical and legal barriers; the specific ethical barriers on Discord could be present on a platform without the legal barriers we encountered and vice versa.

Ethical Barriers

As on Telegram, entering servers on Discord may require a level of deception when answering questionnaires, which are very common for vetting newcomers attempting to enter dedicated communities. Providing deceptive answers to such verification questionnaires is particularly problematic in the case of smaller servers in which users may have reasonable expectations around privacy.

When creating research accounts and subsequently conducting ethnographic research, ISD always tries to minimise deceptive practices. Profile pictures or information about gender or interests are only provided if necessary, and this must be justified on a case-by-case basis. Additionally, account names must be pseudonyms; this is in order to protect researchers but also prevent any harm to third parties. The pseudonyms (or other information provided in line with the principles outlined above) used in accounts should not be identifiable with any real person.

Without passing the verification process, users are confined to a waiting area. The extent of information about the server that is already visible in the waiting area varies significantly. On most servers, there is a questionnaire to verify new users (only 8 out of 47 servers we assessed did not have a verification process). The questionnaires usually consist of a handful of questions. After reviewing these questionnaires, ISD grouped them together into several categories, based on how deceptive and ideological responses would have to be.

- **Level 1:** Questions that require no information about users' political and/or religious views.
- **Level 2:** Questions that ask about users' identity but do not ask anything that is particularly contentious (e.g. 'What is your religion? or 'What continent do you live on?').
- **Level 3:** Questions that ask about users' identity, which additionally require answers that expand on users' religious, political or ideological views (e.g. 'Please outline your political views?' or 'Are you actively practising your religion?').

- **Level 4:** Specific and contentious questions about users' identity, which require answers that demonstrate religious, political or ideological agreement (e.g. 'Do you support suicide bombings?' or 'Do you agree to hate all sin and degeneracy?').

ISD researchers generally answered questions that fell under levels 1 or 2 (21 servers). In all but one case, ISD was permitted to join the server. In the case where access was not granted, server admins asked further questions about political and religious beliefs that would have required unjustifiable levels of deception to answer.

A related ethical issue is that questionnaires often ask users if they agree with the rules of the server. The rules may be highly contentious themselves, including ones that, for example, ban certain religions or political movements or advocate for gender segregation. By agreeing to these rules, respondents may be perceived to be endorsing them.

Another ethical issue that arises in smaller servers is clarity as to where the threshold for reasonable expectations around privacy begins. This is not an issue that is restricted to Discord: WhatsApp groups are limited to 256 members; Signal allows for insecure messaging groups with up to ten members and groups of up to one thousand; Facebook profiles can have up to 5,000 friends. These online contexts may all be considered private (though this depends on settings in the case of Facebook). This indicates that the dividing line between public and private spaces can be somewhat arbitrary as there is no agreed upon numerical threshold at which a private space becomes a public one. The size of Discord servers varies significantly, with the largest non-political Discord servers having memberships in the hundreds of thousands. More frequent, however, are smaller servers with hundreds or thousands of members. Most of the religious extremist servers identified by ISD researchers had just hundreds or tens of members. Consequently, while there is no objective standard to divide between clearly private and public communications channels, including the number of users or participants, it is evident that many of these extremist communities on Discord operate in a grey area in terms of what could reasonably be expected to be a private space online.

Legal Barriers

Discord's API client enables users to connect to a server and collect channel messages live, as well as collect historical messages. In order to connect to a server, researchers must identify themselves using one of two options. The first is a bot account, which needs to be manually accepted into a server by an admin (e.g. the server creator or another member given those privileges), who may refuse such access. Additionally, the bot will be clearly identified as such in the user list, which might raise suspicion, especially among communities discussing sensitive or controversial topics, sharing illegal content or conducting illegal activities.

The second possible option is to run a bot behind a regular user account (a so-called self-bot). In this case, researchers join servers as a normal user (e.g. via an invite link), and the bot subsequently impersonates this user. This deceptive behaviour contravenes Discord's TOS.[42] The TOS also link to the Developer Policy which prohibits scraping – the bulk collection of data from a Discord server ('You may not use the APIs in any way to scrape any Discord Data').[43] To conduct ethnographic research on Discord (i.e. conducting research by acting as an observer or participant in a group) researchers must create a user account. When signing up to Discord, users agree to abide by the platform's TOS. Discord therefore presents legal issues for research as deceptively gathering data from channels via the API would appear to break contract law, thereby potentially risking legal action from the platform.

Analysis of Discord Communities: Key Findings

The fragmentation, ethical and legal barriers outlined in the previous sections combined to limit the scope of our research. Our findings therefore rely on qualitative and non-systematic insights from servers where ethical considerations did not outright prevent us from joining in the first place. The findings detailed in the following sections nevertheless point towards a significant presence of hateful, anti-democratic and violent content within these communities.

Overview of the Servers

The following sections outline the findings of our ethnographic research into Catholic and Islamist extremist communities on Discord. We start by providing a general overview of the size, nature and tone of these communities. We then detail examples of the hateful and violent rhetoric we identified.

In late August 2022, ISD researchers identified support for Catholic extremist views in 31 servers on Discord with a total of 9,585 followers and support for Islamist extremism in 16 Discord servers with a cumulative 4,757 followers. It should be restated that even though community-building is one of the main purposes of Discord, there is significant diversity within servers. It is therefore more accurate to say that our analysis focused on Catholic integralism and Islamist extremism activity within Discord servers where other topics are also discussed rather than servers that are exclusively focused only on Catholic integralism or Islamist extremism.

Despite this, server rules often state that comments may not violate religious teachings, offend believers or attack religious authorities, and that users may be expelled for breaking these rules. Additionally, irony and gamification (i.e. the application of video game elements to facilitate radicalisation, promote extremist views or design extremist or terrorist content) are prominent in most servers. It is often unclear whether users actually support authoritarian religious states or movements or merely find them aesthetically interesting or 'cool'.[44]

Users in both Catholic and Islamist extremist servers often outline their vision for an ideal society. Sometimes admins will also describe it in the server rules. In line with ISD's definition of extremism (see Glossary), these utopian visions generally claim the superiority and dominance of one identity-based 'in-group' (Catholic or Muslim) over all 'out-groups', thereby advancing a world view that appears incompatible with pluralism and universal human rights. The exact form the ideal Catholic state would take unclear; researchers identified users supporting left-wing anti-capitalist integralists, fascists, monarchists and Christian nationalists, among others.

Figure 12: Content and memes shared in Catholic and Islamist Discord servers expressing opposition to human rights and trans rights, as well as advocating for a totalitarian religious state.

Among Islamist extremists, calls to re-establish a totalitarian modern version of the Caliphate that would unify all Muslims are also common. In this Caliphate, all aspects of life, including political, scientific and spiritual matters, would be subjugated to Islamist extremist interpretations of Islam. This includes a strict legalist interpretation of the Sharia, comprising corporal punishments and limits on freedom of belief, freedom of expression and the rights of sexual and religious minorities. Sunni Islamist extremists in the analysed Discord servers reserved particular scorn for Shia Muslims, Sufis, atheists and liberal Muslims.

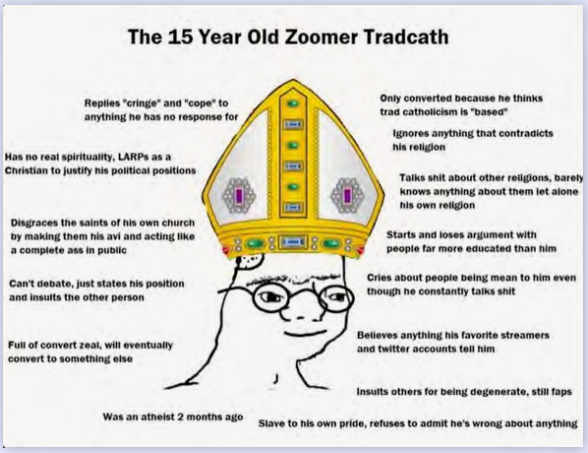


Figure 13: Ironic memes that make fun of clichés relating to Catholic extremist online subcultures.

Hateful Rhetoric

The following section outlines different types of hateful and offensive content towards particular communities that was identified in various servers. While the positive vision for society can be hotly debated, there are recurring out-groups and political enemies. Distancing oneself appears to be critical to the formation of an in-group identity and the affirmation of what is considered 'normal' or 'based.'[iv]

Some server admins appear aware that the ideological leanings of their communities could come into conflict with Discord's policies around hate speech or the promotion of violent extremist content; for example, two Catholic servers asked members in their rules to try not to target protected groups or BAME communities ('Black, Asian and Minority Ethnic', likely a server based in the UK where this acronym has been most commonly used) in order to avoid violating Discord's TOS.[v]

Among both Catholic and Islamist extremist users on Discord, hate against LGBTQ communities serves as a unifying factor. Despite the servers' ideological divergence on many issues, there appears to be a near-consensus on the rejection of homosexuality and trans rights. During discussions in Catholic and Islamist extremist servers, users generally articulated what they perceive to be the normative position of their religions.

iv 'Based' is a slang term that originally meant being addicted to crack cocaine. It was first re-appropriated by rapper Lil B to mean 'behaving authentically' and then by the online far-right to describe things in line with their anti-progressive values.

v Discord's Community Guidelines state: 'Do not organize, promote, or participate in hate speech or hateful conduct. It's unacceptable to attack a person or a community based on attributes such as their race, ethnicity, caste, national origin, sex, gender identity, gender presentation, sexual orientation, religious affiliation, age, serious illness, disabilities, or other protected classifications.' 'Discord Community Guidelines', Discord, February 2023, https://discord.com/guidelines.

Figure 14: Anti-LGBTQ comments and memes celebrating the death penalty for homosexuals by throwing them off a roof and mocking the idea that Islamic scripture is compatible with LGBTQ rights.

Antisemitism similarly unites Catholic and Islamist extremist servers on Discord. Both often emphasise religiously grounded antisemitism that draws on persistent, historical forms of antisemitism. While antisemitic narratives among Islamist extremists are at times loosely connected to hatred of Israel, most of the antisemitic content within both Catholic and Islamist extremists on Discord is unconnected to Israel, instead portraying Jews as Satanists (or even controlling Satan), demons, greedy or sub-human. Other content (see Figure 15) claims that Judaism is not just demonic but also intertwined with advocacy for trans rights; this is a popular talking point in far-right circles, one which argues that Jews are trying to weaken patriarchal societies by undermining binary gender norms.

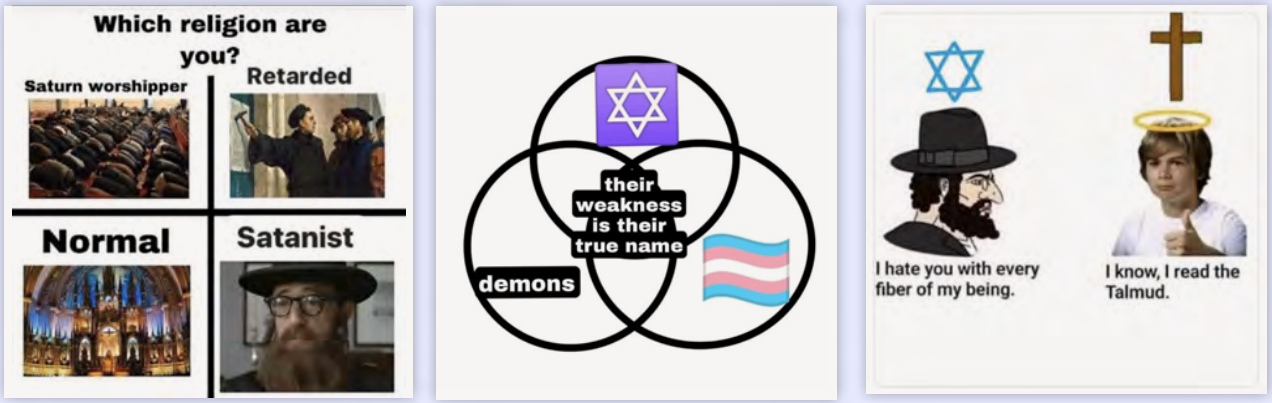


Figure 15: Memes attacking Jews, Protestants, Muslims and Trans People.

Opposition to feminism is another unifying factor both within and between Catholic and Islamist extremist servers on Discord. Users often mock feminists or complain about feminists being 'triggered' by their views on gender norms. One common angle to attack feminism is claiming that traditional gender and family norms would in fact make women happier while feminist values produce old, childless, lonely and regretful women.

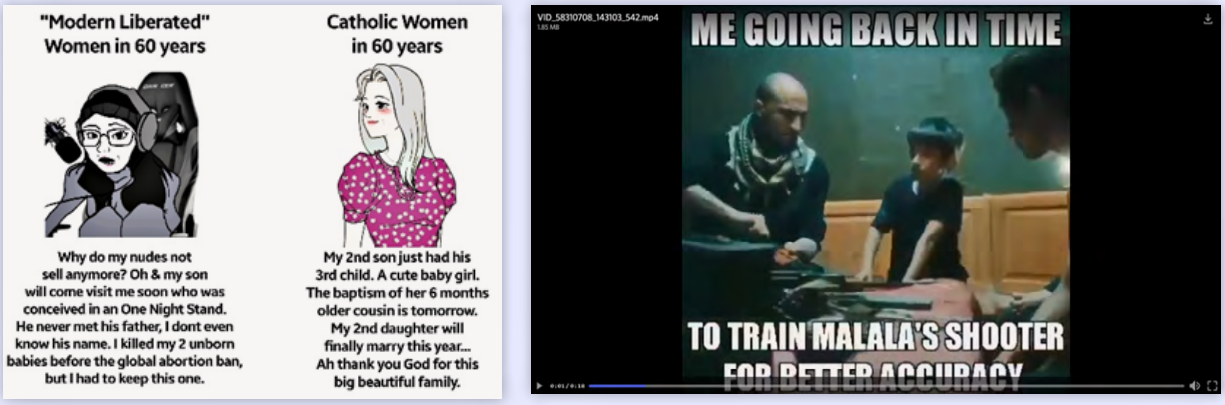


Figure 16: Meme mocking women for not conforming to traditionalist gender roles and screenshot of a video fantasising about killing women's rights advocate Malala Yousafzai.

Another critique of feminism and women's rights that is more frequently found among Catholic extremist servers is opposition to abortion (while there is some diversity of views within Islam as well as among Islamists, opposition to abortion under all circumstances is relatively rare). Users in Catholic servers generally praised the overturning of Roe v. Wade in June 2022, with a left-wing integralist server even denouncing the right to abortion as 'individualist and intertwined with liberalism', arguing that this was why Nikita Khrushchev had legalised abortion as part of his 'de-Stalinisation' reforms in the 1950s. In some Islamist extremist servers supportive of the Taliban, users also expressed opposition to women's rights to education. One video dreamed of travelling back in time in order to ensure that Malala Yousafzai's assassins were trained well enough to actually kill her.[vi]

Violent Rhetoric

When classifying violent content, it is worth considering that users may take stances that fall somewhere on a continuum rather than treating support for violence as a binary variable. This may include an implied need for violence; celebrations of or support for known violent offenders or groups; calls to join violent movements; unspecific calls for violence; or even instructional material and information that could leave individuals or groups exposed to violence. While there were few specific calls for violence, researchers identified a range of unspecific calls for violence, celebrations of violence and posts that implied a need for violence, especially within Islamist extremist Discord servers; for example, users expressed hope that God would allow the Mujahideen to kill homosexuals and their supporters or to initiate a serious insurgency against the Pakistani state. Some memes portrayed the killings of Westerners or the hanging of Hindus (depicted as Pepe the Frog characters named Pajeet, a derogatory online-term used to refer to Indians).

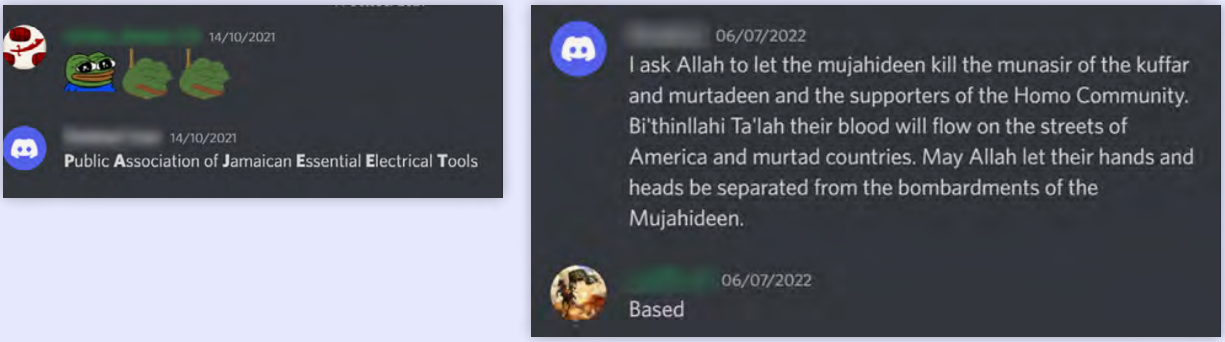


Figure 17: Comments calling for the execution of Indians (symbolised through the name Pajeet) and calling on God to kill homosexuals.

vi Malala Yousafzai is a Pakistani human rights advocate and Nobel Peace Prize Laureate who was shot in 2012 by the Taliban for her activism for the right of every child, specifically including girls, to receive an education.

Within Islamist extremist servers, users similarly expressed support for extremist and terrorist organisations, including ISIS, al-Qaeda, the Taliban and Hamas. This included the sharing of official ISIS and Taliban content (including execution videos) and the creation of their own memes and video game graphics that celebrate extremist or terrorist groups. Interestingly, pro-ISIS content and comments appear to be particularly controversial and tend to trigger more critical reactions than statements supportive of other Islamist extremist movements; for example, the Taliban seem to enjoy more support, with a wide variety of memes celebrating their successful re-establishment of the Islamic Emirate of Afghanistan in September 2021.

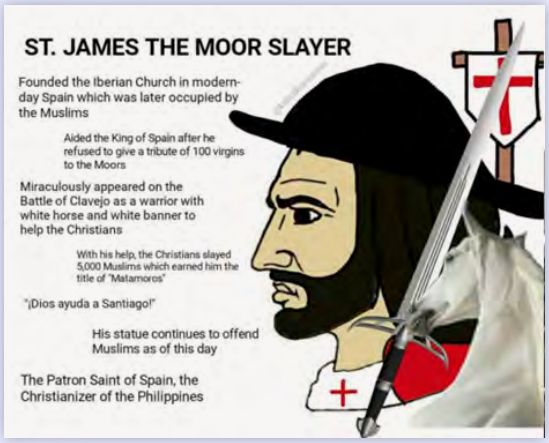
Figure 18: Posts expressing support for the Taliban and ridiculing the idea that religion could be undermined by education by pointing out that the late al-Qaeda leaders Osama bin Laden and Ayman al-Zawahiri were trained as engineers and heart surgeons respectively.

It should be noted that this does not indicate a coherent ideology within the servers analysed. Firstly, the different extremist groups are often opposed to each other; for example, the Taliban, al-Qaeda and Hamas are all opposed to ISIS. In the case of Afghanistan, for example, the Taliban and ISIS-K (the Afghan branch of ISIS) are engaged in military combat with each other. Secondly, as previously noted, there is often a diversity of ideological views within servers, and disagreement between users is commonplace.

Figure 19: Memes and video game graphics expressing support for the Taliban.

Violent rhetoric was less frequent within the Catholic integralist servers analysed. At times, however, users and admins implied that there was a need for violence to fight 'degeneracy and the modern society'. In one instance, a user shared a quote by Charleston church shooter Dylann Roof that asks readers to start fighting for the 'white race' rather than

for alleged Jewish interests (see the previous section on Hateful Rhetoric). Lastly, there were some memes shared that centred around the mythical figure of St James the Moor Slayer and celebrated the killings of Muslims.



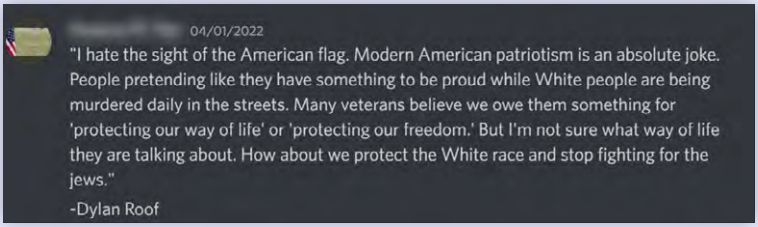


Figure 20: Content centred around the mythical figure of St James the Moor Slayer, celebrating the killing of Muslims, and a quote by Charleston church shooter Dylann Roof.

Findings and Recommendations

Our analysis of Catholic and Islamist extremist Discord servers suggests that there is a small but significant presence of harmful communities that share, at times, very extreme, hateful, anti-democratic and violent content. Due to fragmentation barriers, ethical concerns and legal risks, our research findings remain qualitative and non-systematic, likely only showing a fraction of the true scale of similar activity on the platform.

The ethical barriers are understandable in the context of privacy and human rights concerns, but policymakers, platforms and human rights advocates should work together to develop more straightforward and applicable definitions of public and private spaces which balance these concerns against public interest research on extremism and other harmful online activity.

We contend that the legal risks around potential breaches of contract law presented by Discord's TOS prevent legitimate public interest research of harmful communities. It would be desirable if platforms, including Discord, change their TOS to allow systematic public interest research rather than effectively insulating themselves from public scrutiny. Preferably, the government would pass legislation to protect researchers from legal risks when conducting public interest research of harmful communities operating in public spaces online.

Lastly, Discord should cooperate with Disboard to 1) improve search functionalities and 2) proactively identify and moderate servers that contain illegal activity or Discord's TOS by using Disboard's tagging function.

Case Study 3: **Odysee**

Key Findings

- Odysee is a primarily audiovisual platform built on blockchain technology. We therefore investigated Odysee as a potential example of technological barriers.
- We found that the blockchain-based nature of Odysee presented fewer insurmountable barriers than we might have expected. Indeed, the public visibility of transactions using the LBC cryptocurrency provided additional data points. However, it did require technical expertise and substantial work investigating and amalgamating different tools. The audiovisual nature of the platform also presented challenges as researchers needed to manually inspect content.
- As Odysee currently offers less moderation than more mainstream audiovisual platforms, alongside potential monetisation of content and options for importing videos from YouTube, it could become a popular site for extremist or conspiracy-related audiovisual content.

For the case study looking at French-speaking extremist communities, ISD undertook research into neo-fascist and royalist communities (as well as a cluster of Catholic fundamentalists who link to both) on Odysee, a platform which primarily presents technological barriers. As decentralised and/or blockchain-based platforms like Odysee are relatively unexplored territory in online research, we have tried to explore whether systematic search methods are applicable, what data becomes available and which additional barriers arise during the process.

At the outset of this research project, it was unclear what types of data would be available from Odysee via the LBRY API (see below) and whether the decentralised nature of the platform might lead to issues in synchronising data points. As it turns out, LBRY gives data access to wallet details, block details, transactions, and channel and video details. However, technological barriers do still exist; for example, it is not immediately apparent whether a video was imported from another platform, such as YouTube, or originally posted on Odysee, and there are only imperfect methods available to determine this.

While conducting the research, we identified additional legal barriers to researching Odysee. The platform's TOS place restrictions on the gathering of personal data but do not provide a clear and actionable definition of what this constitutes in this context. Since the data in question is publicly accessible and visible without logging into an Odysee account, ethical considerations around privacy are limited. It is, however, unclear if collecting what Odysee defines as personal data would break contract law.

These barriers combine to limit the data points that can be gathered for this case study and the way in which they can then be presented. For the purposes of this research, we decided to collect data that was posted publicly, such as videos and comments, and aggregated it in user categories. The findings point towards a presence of extreme anti-democratic ideologies within French-speaking communities on the platform. This includes material denying the Holocaust and venerating National Socialism, both of which are potentially a criminal offences in France.

The following sections outline the background and functionality of Odysee, the research approaches that we scoped and used, the limitations encountered and the findings of our analysis.

Platform Overview

Odysee is a video platform that positions itself as an alternative to YouTube. In December 2020, the platform claimed to have 8.7 million users.[45] While the platform is not exclusively used by extremists and conspiracy theorists, it has become popular among those banned by larger video platforms for violating their TOS because Odysee wants to present itself as taking a more lenient approach to content moderation than more established ones;[46] for example, when questioned why they allowed the Russian broadcaster RT (now sanctioned in the EU) to continue using their service, Odysee released a statement that they 'want to allow competing voices in journalism'.[47] However, a recent report by ISD Germany found that Odysee was actually blocking too much content and restricting whole accounts over a few videos that violated guidelines, while some illegal content remained visible for German IP addresses.[48] Odysee has distanced itself from the label alt-tech, which connotes ideological proximity to the far-right, preferring the name 'new tech'.[49]

Odysee is the successor to a video platform called LBRY.tv and operates on the LBRY network, a blockchain-based file-sharing and payment protocol owned by LBRY Inc. According to its own website, LBRY wants to do for publishing online content 'what Bitcoin did to money', namely create a decentralised alternative to established platforms, thereby promising users more control over their creations.[50] As the LBRY network is a blockchain protocol, this theoretically means that content cannot be moderated by a central authority, although Odysee does moderate some content in practice. LBRY also has its own currency called LBRY Coin (LBC) that users can earn by uploading videos and receiving interactions with their content. Importantly for researchers, because LBRY is a public blockchain, these financial transactions are visible; we expand on how this can be used below.

The CEO of LBRY Inc., Jeremy Kauffman, was previously a New Hampshire Senate candidate for the Libertarian Party in the US in 2022.[51] LBRY's affiliation with libertarianism likely contributes to the perception among users that the platform pursues only limited moderation.[52] In October 2021, Odysee announced that it would develop its business independently from LBRY but still use its protocol and cryptocurrency. Odysee has a separate CEO and runs via a LBRY subsidiary called Odysee Holdings Inc.[53] The goal of the new CEO, former TikTok employee Julian Chandra, is apparently to move Odysee from the libertarian space to a more mainstream audience in order to make it profitable.[54]

Key Functionalities

How Odysee Hosts Content

Odysee mostly serves to upload videos but can also be used to distribute other file formats like text or audio. In order to upload or engage with content, users have to create an account and thereby agree to Odysee's TOS.[55] However, it is possible that there are videos with no attribution on the platform (i.e. there can be videos where the uploader is anonymous and the original uploader's account is no longer accessible). Users can also import their videos from YouTube by syncing their channels while signing up for Odysee.

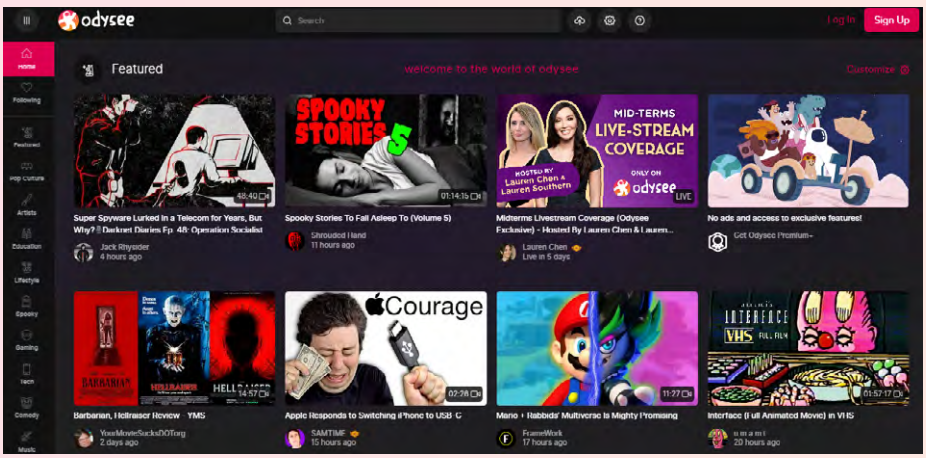


Figure 21: Odysee's homepage (03/11/2022).

While the front page of Odysee offers a wide variety of content, both political and non-political, it sometimes features controversial political figures, supporting the idea that fringe extremist ideologies and disinformation are accepted on the platform. The Odysee sidebar features different content categories, many of which relate to hobbies and lifestyle. However, the 'News & Politics' section includes channels such as RT, Sputnik and The Alex Jones Channel, which fuels the assumption that Odysee does not moderate demonstrably and intentionally false content.

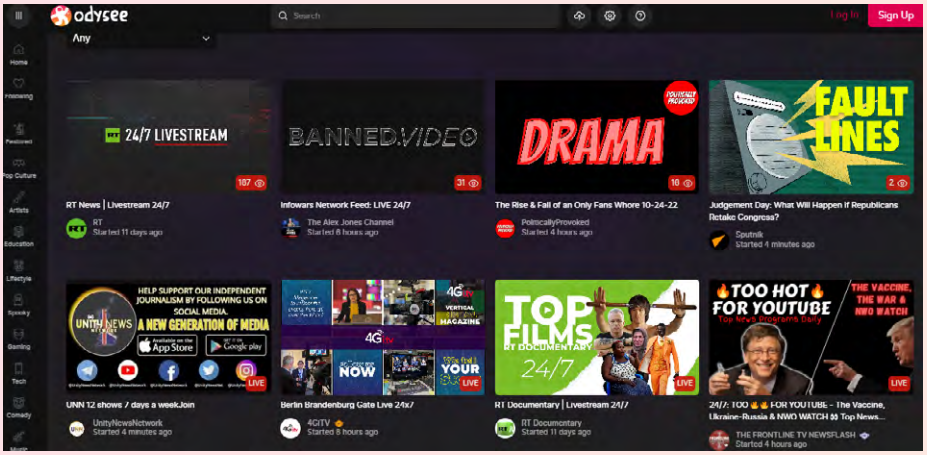


Figure 22: Odysee's 'News & Politics' section (08/11/2022).

Monetisation

Odysee allows creators to monetise their content with the aforementioned LBC cryptocurrency. LBC can be traded for official currencies like the US dollar on cryptocurrency trading platforms. In late October 2022, 1 LBC was worth only $0.02 USD. Transactions on the LBRY blockchain are public, and it is therefore possible to determine how much a user has made in LBC on Odysee. Every account comes with an LBC wallet attached, and every video has a transaction ID. Odysee's use of blockchain – to more easily transfer currency from a wallet to a creator on the same platform – lowers the threshold for monetisation than on non-cryptocurrency-based platforms like YouTube; these platforms require more formal partnerships with creators to enable monetisation, which also allows the platforms to rescind creators' ability to monetise content if it is deemed to not be advertiser-friendly. As a result, Odysee's use of the blockchain potentially makes the platform attractive to those barred from monetising their content on other platforms.
Although Odysee did not originally run ads, it announced in October 2021 that it would start running commercials over videos, with the promise to take a smaller cut from the revenue than other companies.[56] The platform also offers a premium plan with ad-free viewing and various other features.

Harmful Activity on Odysee

Many self-proclaimed 'censorship-free' platforms tend to attract users who were banned from others for posting extremist content or disinformation. Platforms that espouse this view, such as Odysee, may choose not to remove these users due to ideological conviction; because they can gain new users when fans of the banned individuals join them on the new hosting platform; or because they simply lack the resources or expertise to employ a more comprehensive approach to moderation. Odysee has received negative attention in the French media as a preferred platform for conspiracy theorists, who were barred from YouTube for posting false claims about the Covid-19 pandemic. The French-language disinformation 'documentary' Hold-Up appeared on Odysee after being removed from YouTube. Several leading far-right figures like politician Florian Philippot as well as disinformation outlets like France-Soir have established channels on Odysee. Since monetisation is available to any user, Odysee presents a possible risk as a platform for funding anti-democratic activities even though the current exchange rate for LBC is fairly low.

Researching Odysee

Research Methodology

While it does not appear that there is an API specifically for Odysee, the data required to conduct this research on Odysee could be accessed via APIs built for LBRY. We investigated two of the tools that LBRY provides to interact with the LBRY blockchain and network: Chainquery and LBRY SDK.

- **Chainquery:** This is a tool to query transaction data on the blockchain. Although easy for a researcher to use, Chainquery requires access to a local copy of the LBRY blockchain, which necessitates some difficult set-up and maintenance, including more powerful hardware. There are also additional security and ethical concerns involved in joining the LBRY network as this would require responding to other users querying the blockchain, thereby actively participating in the network while also researching it.
- **LBRY SDK:** This tool provides resources for developers to create their own applications to interact with the blockchain. In addition to querying, a researcher can also interact with the data (e.g. find all the parts that a video has been split into and download them). With LBRY SDK, users can connect to the network and access information about content and transactions as well as related metadata. While creating data access requests for the LBRY SDK API can be more time consuming, the setup is easier and maintenance requirements are less significant.

For this project, we chose the latter approach to integrate Odysee into Method52. Neither of the LBRY tools provide access to the comments posted on Odysee. This is handled by a separate system, Commentron, which is managed directly by the Odysee team. The system also provides an API, which functions in a similar way to LBRY SDK. In order to have access to video comments for research, it was necessary to integrate this API into Method52 as well.
Researchers can also analyse Odysee in a similar way to other video platforms. That includes manually reviewing video content, investigating the backgrounds of channel owners and collecting statistics provided to all users by the platform, such as the number of subscribers and uploads. However, manually reviewing video content can become laborious with any larger sample of content and accounts. It also does not usually allow for more detailed systematic analysis, such as investigating spikes in activity around certain time periods.

Manual research of financial transactions by relevant actors allows analysts to identify crypto-wallets and their balances. However, this is again very time consuming when assessing a large seed list of accounts. Furthermore, there are some limitations with analysing the credit balances of multiple wallets. Users can own more than one wallet, and therefore it can be difficult to understand the true financial earnings of any single user on the platform. Additionally, if monitored channels are sending each other credits, these LBC could be double-counted when only comparing the opening and closing balances of a wallet.
To automate the process of collecting LBC earnings, CASM built a component for Method52 that collects a video's details, including wallet IDs and support amounts (i.e. how much LBC users earned with their videos). As well as enabling larger scale data collection via automation, another benefit of this approach is that wallet IDs can be attributed to accounts with greater certainty. The LBRY Block Explorervii does not make it clear which wallet ID belongs to which channel, and researchers using a manual approach must therefore infer relationships between users and wallets by analysing the individual transaction IDs of videos. This is not only time consuming but also does not conclusively prove which wallet IDs belong to which channel.

Technological Barriers

Blockchain is still an emerging technology and there is limited systematic, data-driven research on blockchain-based social media platforms. As a result, there can be unexpected results when using novel data collection methods, and the decentralised nature of the technology might lead to distortions in the collected data (as ISD found when researching another decentralised platform, PeerTube). Extremists' use of blockchain technologies has increasingly concerned researchers in other contexts, for example, after ISIS supporters started using NFTs to create non-erasable terror propaganda.[57] Commentators, however, have largely focused on the use of cryptocurrency rather than the use of blockchain-based platforms for propaganda.

vii The LBRY Block Explorer is a publicly available online tool that enables users to search for both real-time and historical information about a blockchain, including data related to blocks, transactions or addresses.

While Odysee also provides opportunities to upload text files (e.g. PDFs), it is predominantly a video platform. Audiovisual content presents a particular challenge as it is more difficult and laborious to systematically analyse a large amount of this type of content. Unless subtitled are available, text-based analysis is usually limited to video titles, descriptions and comments. However, subtitles (especially those that are automatically generated) might still cause issues due to unclear audio or the use of multiple languages.

There were some challenges during the data collection, which related to the configuration of Odysee as a platform. The initial collection hit a limit once 1,000 videos were collected from a channel. This appears to be either a feature or unintentional bug in the design of Odysee where only 1,000 videos are displayed on the channel. Older videos are not removed but are not visible on the channel. The data collection therefore had to be adjusted to access these videos that were not on display.

The software ecosystem around Odysee is also disjointed and overlapping when compared to many other online platforms that offer a more coherent set of tools for data access. While there are a variety of official tools available to interact with Odysee, including LBRY SDK and Chainquery, looking through them takes time and significant technical expertise. These tools do not have the same level of documentation and so implementing them can take considerable time. Both tools offer many of the same functionalities and access to data but also differ in other ways, making it difficult to determine which is most appropriate for different purposes; for example, in order to have access to both video and comment data, it was necessary to integrate two different APIs since these systems are handled separately by the platform.

Overall, the technical barriers to research posed by Odysee were not as prohibitive as we initially expected; ISD's previous research on the platform had not attempted to employ a more technological approach. However, while we were able to collect data from the platform using technical means, the barriers to entry for this type of research were relatively high, requiring significant technical expertise and time for researchers to familiarise themselves – and experiment – with the range of options available.

Ethical and Legal Barriers

Since most content on Odysee is publicly accessible even without logging into an account, there are limited privacy risks. There is a reasonable assumption that users are aware their content can be viewed by anyone that finds it. The LBRY Block Explorer that registers users' wallets is also publicly viewable and shows wallet balances all without having to register or log in to Odysee or LBRY accounts.

The main ethical and legal issues are related to Odysee's TOS, which prohibit 'harvest[ing] or otherwise collect[ing] information about users, including email addresses, without their consent'.[58] The phrasing of the TOS is unclear as to which types of the available data should be considered personal and therefore should not be collected. Email addresses are not public on Odysee and cannot be accessed through the available official tools; this raises questions around whether it is legal and ethical to collect the various types of data that are available, such as uploads and views on public channel activity, provided that the corresponding users are anonymised.
For the purposes of this project, ISD and CASM decided to collect data via LBRY's official API while taking into account the TOS outlined by Odysee itself. Once it was determined which types of data could be collected through the available API, we assessed which should be considered personal data. For this project, we decided to collect data that was posted publicly, such as videos and comments, and aggregate it into categories of users. This approach ensured that researchers could not identify single individuals within the data, easing the potential ethical and legal concerns posed by Odysee's TOS regarding the collection of personal data. However, this approach did prevent the analysis of certain data points that would have enabled a more detailed understanding of user behaviour on the platform.

Analysis of Odysee Communities: Key Findings

There is a broad but ideologically fragmented ecosystem of French-language far-right and disinformation communities on Odysee. Since this ecosystem is too broad and ideologically diffuse to analyse in full, we instead focused on three specific subgroups on the French far-right that have ideological connections to each other, although their ideas are not fully congruent. These ideological factions are far-right royalists, neo-fascists and Catholic fundamentalists. Far-right royalists were chosen due to their historic significance in driving ultra-nationalism and their opposition to the current French political order. Neo-fascists were selected as they frequently post content that is illegal, such as the negation of crimes against humanity committed during WWII. Catholic fundamentalists were included as it was observed they often present an ideological bridge between the two former communities by embracing both ethnonationalism and opposition to secularism while also proliferating conspiracy theories and disinformation, such as anti-vax content. Since all these groups are prone to post conspiracy theories about shadowy elites and often engage in discourse that disparages ethnic or religious minorities, Odysee's reputation for lax moderation may help to attract them to the platform.

Far-right royalists oppose the post-revolutionary political order in France, particularly one of its defining characteristics, the separation of church and state. The more radical branch of French royalists supports a return to 'rule by divine right' and absolutism, where the monarch only answers to God. Such notions are inherently anti-democratic because the ruler cannot be held accountable by any worldly power, such as France's Senate or Parliament, and democratic decisions or authorities can therefore be dismissed. The expression 'extreme right' in the French context was first used for uncompromising royalists who, during the post-Napoleonic Restoration (1815–1830), sat on the right-hand end of the benches.[59]

Neo-fascists do not necessarily tap into Catholic fundamentalism as some reject Christianity in favour of neo-pagan beliefs. However, there is nonetheless a significant overlap between the two, especially when neo-fascists define the people of France as Catholic and of pure French descent and spread conspiracy theories accusing Freemasons and Jews of nefarious activities or hint at Satanic subversions of society.

The Odysee channels that we coded as Catholic fundamentalist focus primarily on religious content (frequently taking the form of conspiracy theories) but often also combine the royalist desire to abolish secularism and the neo-fascist's support for ethnonationalism. They sometimes even promote content denying the Holocaust or other crimes committed during WWII. Ultimately, all these communities reject the enlightenment ideas of universalism and individual rights that inspired the French Revolution and still define the Fifth Republic.

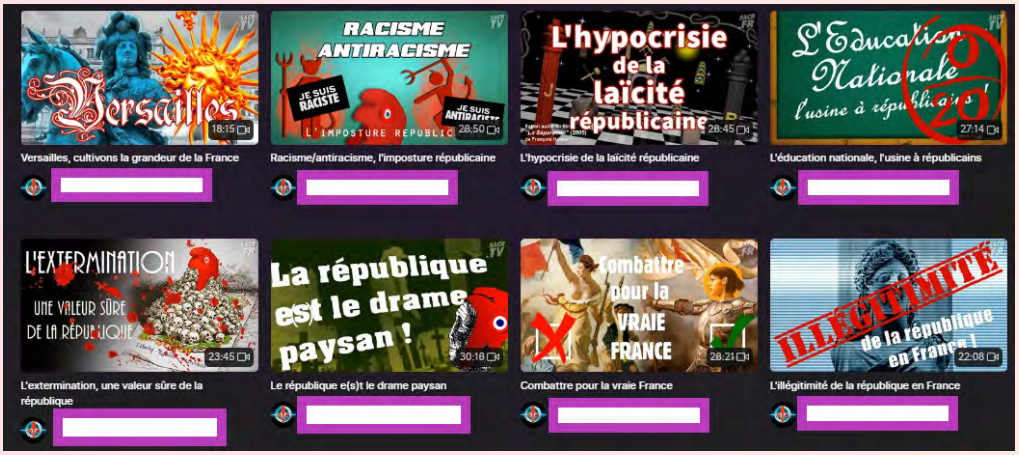


Figure 23: Example content from a far-right royalist Odysee channel.

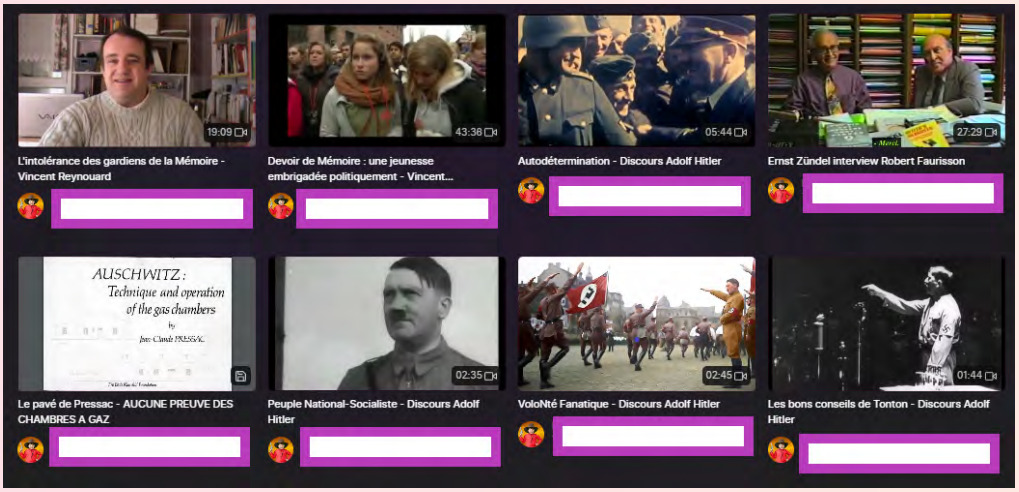


Figure 24: Example content from a neo-fascist Odysee channel.

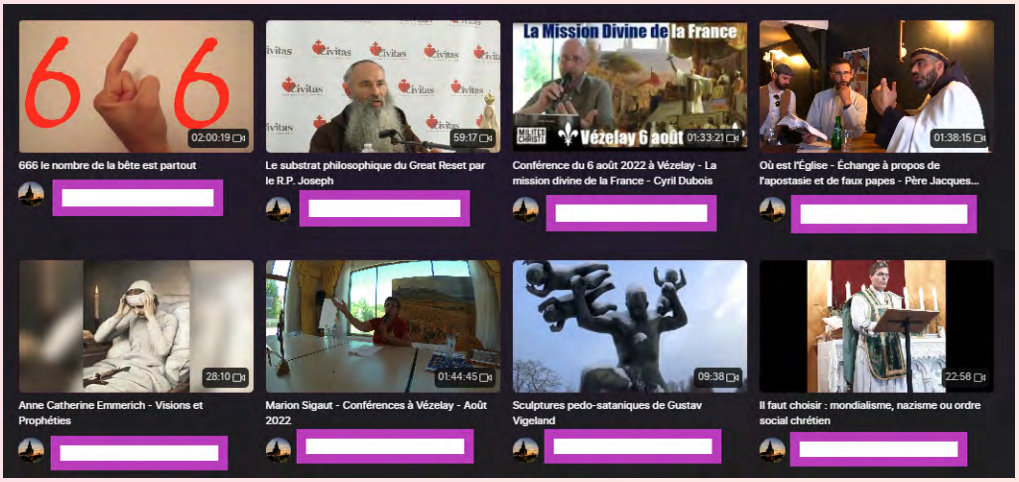


Figure 25: Example content from a Catholic fundamentalist Odysee channel.

Platform Usage and Engagement

Systematic analysis of the content found in these communities poses a significant challenge, even with the relatively limited sample used for this study. Whereas text can be easily analysed using an NLP programme capable of classifying text, audiovisual content needs to be transcribed first. The quality of transcription tools can vary from language to language. Multiple people speaking at the same time or multiple languages being used in the same video usually pose considerable challenges for transcription programmes. Furthermore, visual imagery – including dog-whistles – that might provide important additional context regarding ideological allegiances or the narratives employed can be lost through this approach. For this reason, the following section looks only at data that could be collected using LBRY's API, which consists mostly of channel and video statistics as well as LBC earnings.

Analysing the collected data, covering the period from 1 January 2020 to 15 September 2022, reveals that the first notable spike in activity took place during summer 2020, which is around the time Odysee was launched as a replacement for LBRY.tv. The data indicates that activity from the observed accounts was minimal when operating on LBRY.tv. The initial spikes in activity following Odysee's launch appear to have been driven mostly by uploads from the neo-fascist category. The data for all user categories together shows that since the first rise in uploads in late summer 2020, there have repeatedly been increases in daily upload activity, though daily uploads usually ranged somewhere between ten and fifty videos.

A curious aspect of the data is that much of the upload activity throughout 2021 came from royalist channels, who then became the least active group in 2022. Nevertheless, with 8,690 videos, they uploaded the most content out

of any user category within the observed time frame of just over two and a half years. The second most active user category overall was the neo-fascist channels, who posted 6,035 videos within the observed time frame. Of the three categories, they posted most regularly since Odysee's official launch, with periods of increased activity observable in late summer 2020, autumn 2021 and summer 2022.

The Catholic fundamentalist channels posted only 4,084 videos although they clearly increased their daily upload rate over the observed time frame. While there were only a few uploads from this user category in 2020 and 2021, they became fairly active throughout the course of 2022. Particularly notable is a spike in activity around February and March 2022. Overall, one could infer from this data that the usage of Odysee varies throughout time and by user category. While the royalists appear to have lost interest in the platform, the Catholic fundamentalists are increasingly using it.

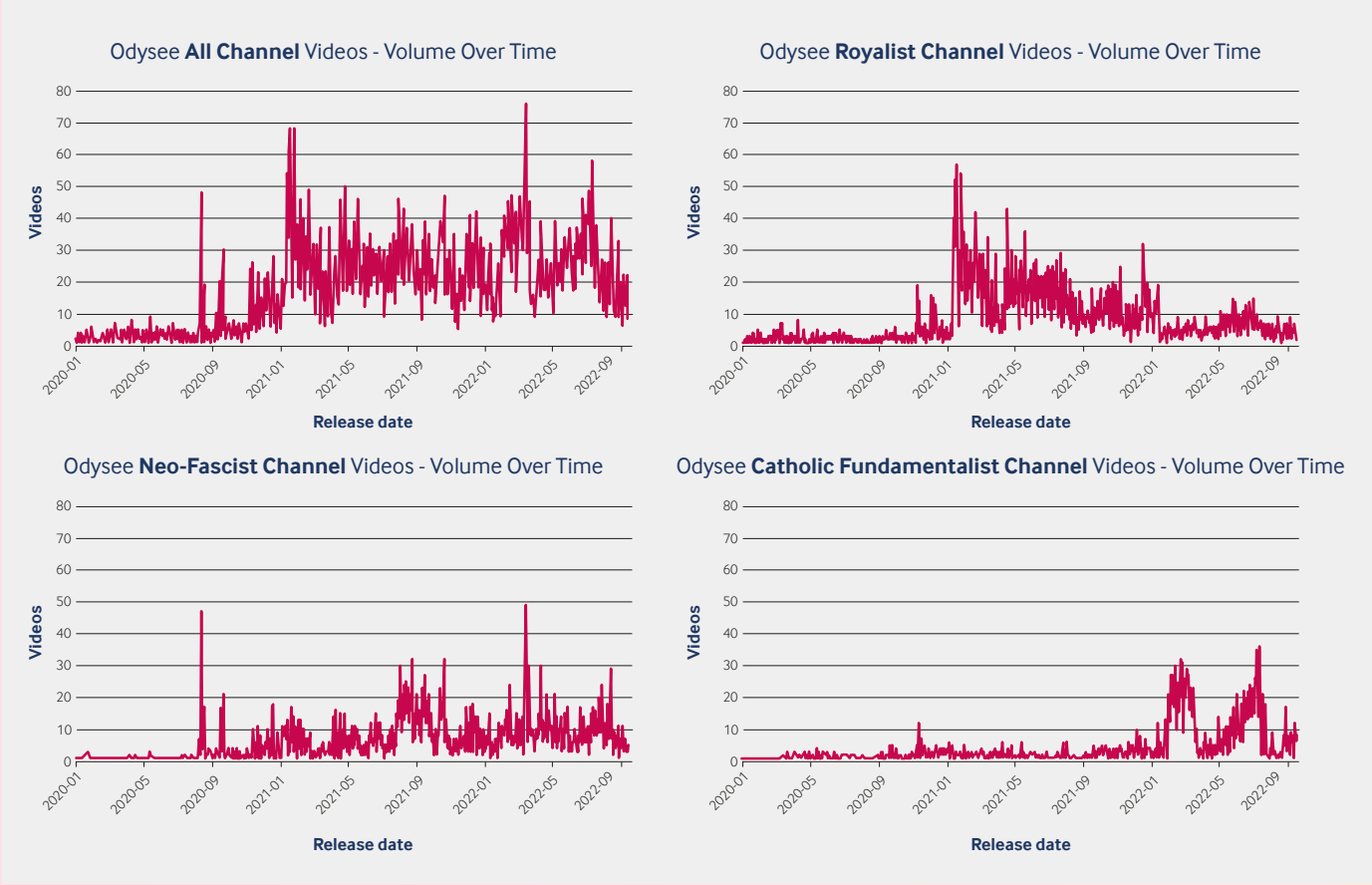


Figure 26: Graphs showing video uploads between 01/01/2020 and 15/09/2022 for all categories and per user category (royalist, neo-fascist, Catholic fundamentalist).

Comments were collected through a separate API and are only shown from mid-2020 onwards. Differences between user categories are also visible here although the average number of comments and replies is rather modest; videos from neo-fascist channels received on average 3.3 comments, royalists 2.4 and Catholic fundamentalists 1.5. This indicates that the upload rate is not necessarily related to how many comments a certain user category will receive.

When examining the volume over time, some correlation with the video uploads is to be expected, though they are not fully congruent. While royalist channels were very active in uploading around the beginning of 2021, these high levels of activity are not reflected in the posting rate of comment. Nevertheless, their declining upload rate was accompanied by a relative reduction in comments. Neo-fascist videos received a fairly steady stream of comments since autumn 2021. Before that, the volume of comments was comparatively low. There are occasional outliers when a particularly large volume of comments is posted in a single day. The closest relation between uploads and comments can be

observed among Catholic fundamentalists who show increased activity in late winter and summer of 2022.

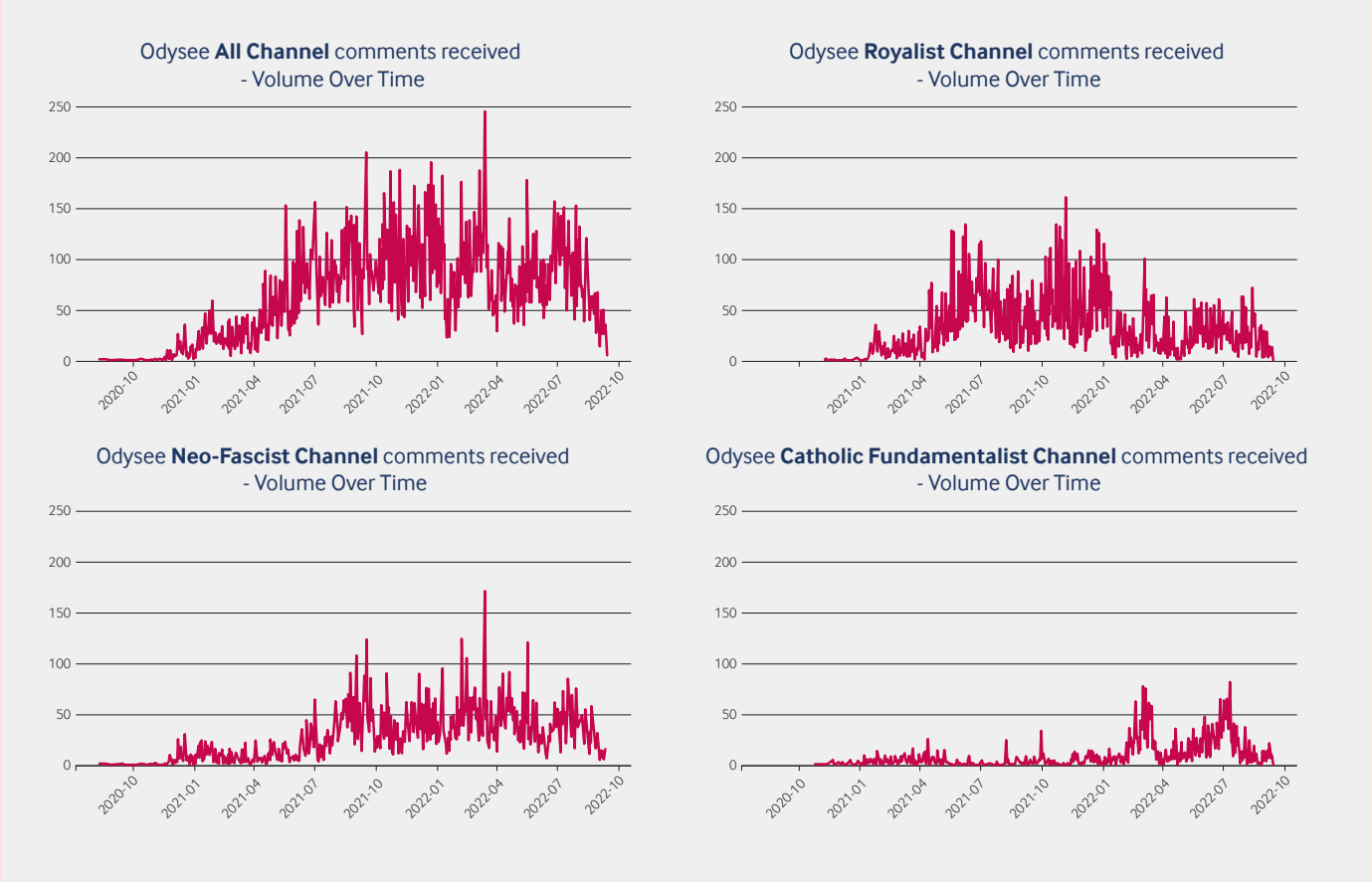


Figure 27: Graphs showing the posting of comments and replies between 13/08/2020 and 15/09/2022 for all categories and per user category (royalist, neo-fascist, Catholic fundamentalist).

Overall, these French far-right communities appear to use Odysee as an alternative to YouTube. Theoretically, the platform allows them to earn money with their content more easily than non-cryptocurrency platforms; however, given the fairly low value of LBC, Odysee's reputation for allowing content that would be removed from other platforms for violating their TOS currently seems to be a more significant factor.

Furthermore, Odysee provides options that many other video-focused platforms lack. Videos can be re-posted and are then shown on a users' own channel page. This function allows researchers to observe ideological proximities and shared ideas among the communities and identify the narratives that serve as bridges between them. In particular, these include antisemitic conspiracy theories and alleged plots to undermine Catholicism and the French people, as well as more global conspiracy theories related to the so-called 'Great Reset' or the 'New World Order'. Odysee also allows users to upload non-video files. Neo-fascists especially use this function to upload documents that allegedly disprove the Holocaust and promote antisemitic conspiracy theories.

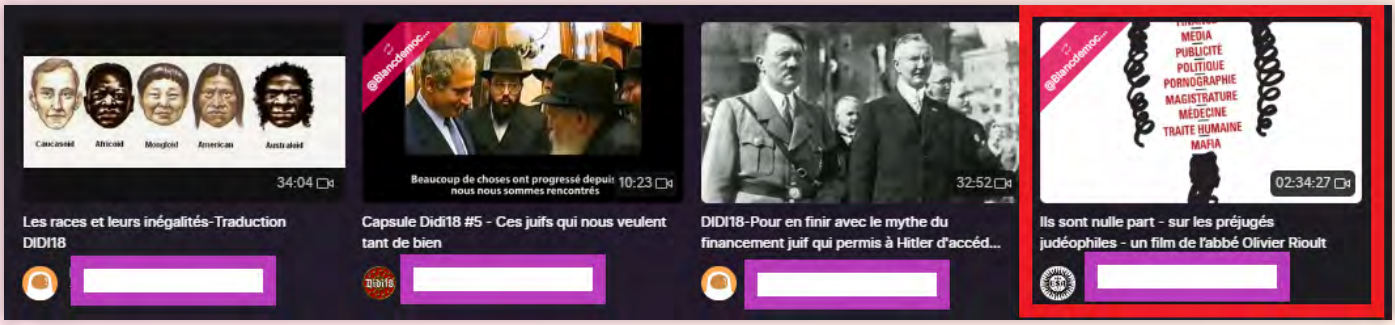


Figure 28: Example of a neo-fascist channel reposting an antisemitic video [marked in red] from a Catholic fundamentalist channel.

Another special feature of Odyssey is that it allows users to directly import content from other video platforms, such as YouTube. This feature had an impact on the data we collected as some videos have upload dates from before Odysee or its predecessor LBRY.tv even existed. It is not immediately clear form our data whether a video was imported from another platform. In order to estimate how many videos were imported from YouTube, all video descriptions were searched for the string 'www.youtube.com/watch?v=' and their volume over time was compiled for analysis. Videos that were imported from YouTube should feature a link to the video's source URL beneath an ellipsis at the bottom of the video description. While this method can provide an approximation for how many videos were imported over time, it is not perfect. Firstly, there might be instances where a video description features the string 'www.youtube.com/watch?v=', but is not related to a file import and might simply advertise another video. Furthermore, the publishing date refers to the upload of the original video to YouTube, not the date when the video was imported to Odysee; this means that the final volume-over-time graphs end up with a mixture of upload dates.

The existence of the import function raises the question of whether the observed actors are attempting to redirect their audience from YouTube to Odysee, which they could perceive as a safer host for their controversial material in case action is taken against their channels by the larger platform. Based on manual research and anecdotal observation, there seem to be two trends that are determined by the type of actor. Firstly, actors which are effectively serving as alternative media outlets that produce their own content, such as podcasts or news commentary, usually pursue a multi-platform strategy, of which Odysee is one of multiple hosting alternatives to more established platforms. These actors often link their profiles on both mainstream platforms and alt-tech platforms like Gettr, Telegram and Rumble, and often use the import function, thereby synching their videos across platforms. It should be noted that these more organised actors usually have more subscribers and views on YouTube than on Odysee. It appears that such outlets mainly use Odysee as one of several backups to keep their content online in case their account on a larger platform is shut down.

The second trend concerns actors who have switched from YouTube to Odysee or have (as far as can be traced) only ever been active on Odysee. These accounts frequently post content that would not only violate the TOS of many platforms but also potentially French law; this includes content denying the Holocaust and glorifying National Socialism. In contrast to the aforementioned trend, these actors have more subscribers on Odysee than they had on YouTube. While on YouTube they had a few dozen and sometimes less than ten subscribers, their followers often number in the thousands on Odysee. These audiences do not appear to be huge, but it should be noted that Odysee is a much smaller platform than YouTube, with the former amassing 23.4 million monthly visits compared to the latter's 33 billion in September 2022, according to SimilarWeb.[60] Channels that feature videos on Odysee's homepage, which can be assumed to show the website's most popular content, may only have around ten thousand subscribers. Those following this trend usually do not import content from their YouTube channels, most of which had been dormant for months when this research was conducted. Instead, these actors started directly uploading to Odysee, which appears to have offered a fresh audience.

Unlike the first group of actors, those behind these Odysee-only accounts typically use pseudonyms. These anonymous accounts often upload footage from other people, for example, interviews with prominent Holocaust deniers or historic footage idolising the Third Reich. One exception is a known neo-Nazi activist who has been previously convicted in France for Holocaust denial but still openly runs his own channel on Odysee. If one examines the data by

the number of YouTube links in the video description, royalists had apparently the most video imports (3,595), Catholic fundamentalists had the second most (1,807) and neo-fascists had the fewest imports (680). One explanation for this divide might be that many royalist channels belong to organised groups and media outlets that pursue a multi-platform strategy. In contrast, neo-fascists might be hesitant to upload their content to YouTube where it is more likely to be moderated. Furthermore, since many of these accounts are anonymous and largely unattributable, they are potentially less interested in promoting themselves via a multi-platform strategy.

In either case, Odysee seems to primarily serve as a way to avoid bans on more strictly moderated platforms for these particular communities. For organisations and media outlets, Odysee appears to serve as a means to protect their content should they be banned from other platforms. For those posting openly extreme content, Odysee has become a primary outlet to spread that content. Although Odysee has a much lower bar for monetisation than YouTube and much of the examined content would likely not be monetised on major platforms, the potential financial benefits are unlikely to be the determining factor for the observed communities in their decision to use the platform. The LBC income made by these channels since the start of our data collection period in January 2020 is negligible. Altogether, the observed channels earned less than $500 USD for almost 20,000 videos in about two and a half years, suggesting that Odysee alone does not (yet) offer a viable business model for these users.

User category	No. of Videos since 01.01.2020	LBC income since 01.01.2020	USD conversion (22.09.2022)
Royalists	8690	4266.733714	$94.66
Neo-fascists	6035	7979.900723	$177.03
Catholic fundamentalists	4084	7736.055632	$171.62

Findings and Recommendations

Odysee has attempted to position itself as an alternative to more mainstream online video platforms, such as YouTube, both in terms of its use of blockchain technology and cryptocurrency payments and its lax moderation standards. Our findings demonstrate that it has proved attractive to some French far-right communities although their presence is still relatively small when compared to other platforms and they do not appear to be generating significant revenues from hosting their content on the platform. Since Odysee's TOS do in fact prohibit the incitement of hate and violence based on ethnicity, religion and nationality (among other characteristics), the presence of overtly racist, antisemitic and fascist material is indicative of a lack of enforcement of these terms.[61] As some of this material is likely illegal in several different jurisdictions, Odysee could come under increasing pressure to remove such content.

Overall the blockchain technology underpinning Odysee did not prove to be as severe a barrier to data access as initially feared. This was due to LBRY providing a variety of existing tools for developers. However, navigating available tools for studying this blockchain-based platform still required technical expertise (which may not be available to many other organisations) and considerable time. Complications arose due to the fragmentation of Odysee's underlying technology, which made it necessary to investigate and integrate multiple tools in order to access the data required for this research. Additionally, not all of these tools were well documented or maintained to the same standards, further increasing the technological expertise and effort required. Smaller platforms like Odysee typically have fewer staff and financial resources than larger tech companies to maintain these services; this can lead to technical challenges and inconsistencies when collecting data for research purposes. Novel blockchain technology can worsen this as the smaller number of experts and tools available may mean researchers are forced to use less well-developed and documented approaches. Where possible, the platform should look to streamline this range of tools and provide more comprehensive, consistent and accessible documentation to accompany them.

The Odysee TOS were also particularly unclear with regards to what constitutes personal data, creating potential legal barriers to public interest research on the platform. This, combined with the availability of a wide variety of different categories and types of data, meant that we took a more conservative approach to our data collection and analysis than would have been technically possible. The platform should clarify its expectations regarding the use of data collected by third parties; this would provide more certainty to researchers and enable users to better understand how their data may be used.

Conclusions and Recommendations

Overall, the research we have conducted for this report illustrates that, while it is possible to conduct digital research on platforms such as Telegram, Discord and Odysee, it can be difficult to do so in a way that is simultaneously systematic, ethical and legal. We were nevertheless able to conduct useful research on each platform, especially with more manual or ethnographic methods; this was complemented with some more systematic, larger-scale data collection and analysis.

However, there were also significant limitations on each platform that prevented the use of more comprehensive research methods and approaches. Ethical considerations around the right to privacy or expectations of privacy prevented us from researching smaller, more harmful Telegram communities. Legal concerns around breaking TOS (and therefore contract law) stopped us from systematically collecting data from Discord servers altogether. Although there were fewer ethical concerns with studying Odysee, given its public nature, a combination of technical complexity and unclear TOS regarding data collection also made more comprehensive analysis challenging.

These technological, ethical and legal, and fragmentation barriers encountered on Telegram, Discord and Odysee are also deeply interconnected and overlapping; for example, legal considerations prevented us from overcoming fragmentation barriers on Discord through systematic search methods, and technological features on Telegram created significant ethical barriers. The complications of integrating multiple tools to access Odysee data were worsened by the platforms' TOS, which are unclear as to whether different tools are providing personal data, thereby creating ethical and legal uncertainty. Ethical and legal barriers therefore often limit researchers' ability to study harmful communities on these platforms in a systematic way even where such approaches may be technically possible.

Implications for Researchers

As our research has demonstrated, platforms can inhibit public interest research in a number of ways, whether deliberately or inadvertently. As noted above, unclear legal terms can create additional burdens or risks for researchers. Technological choices can prevent access to platform data or hosted content or make data collection unnecessarily burdensome from a technical perspective, for example, where platforms' front-end or back-end architectures are fragmented or poorly maintained or the tools available to access data are poorly documented. The design of platforms can also create ethical uncertainty, for example, where online platforms or spaces are denoted as private but are easily and widely accessible to both users and researchers in practice.

Our research during this project required CASM's considerable expertise to explore the technical possibilities on Discord and Odysee in particular; this expertise may not be available to all researchers, particularly in civil society. We also benefited from access to pro-bono legal support that enabled us to more carefully assess platforms' TOS and any resulting legal risks related to data access. This undoubtedly raises the costs and barriers to entry to researching these platforms. Given the significant variety with which these barriers appear and combine on different platforms, each one can present a unique and often complex set of challenges for researchers. To illustrate the scale of the challenge for researchers, in the Phase I report published as part of this project, we identified 81 unique platforms or online services linked to by the English, French and German extremist communities we examined. Without strong mechanisms and incentives for collaboration across the research community, there is a risk that these challenges will be addressed in a disconnected and siloed way, wasting scarce resources and foregoing opportunities to achieve greater economies of scale.

Platforms can also deliberately invoke legal or ethical concerns, such as data protection or users rights to privacy, to stifle external research; in some instances even where users have actively consented to the sharing and use of their data for specific, limited research purposes.[62] Such a response from platforms, in addition to TOS that are unclear or explicitly restrictive to public interest research, can create a perverse and problematic set of incentives for researchers

when it comes to balancing or mitigating different legal and ethical risks. For example, there may be instances where technological means are available to collect data from a platform (e.g. third-party APIs, scrapers or plug-ins for crowdsourced data collection), but their use is prohibited by the platform's TOS. This may encourage researchers to employ deceptive means to connect to a platform, an ethically dubious approach if the platform or users are unaware of or have not consented to data collection.

In such instances, deception may actually serve to reduce legal risks for researchers utilising these methods or tools – if platforms or users are not aware the data collection is taking place, then the likelihood of researchers facing potential legal action is reduced significantly. The use of such methods can also disincentivise researchers from publishing their work or require them to be vague about the methods and tools they have employed when they do. This in turn has a negative impact on the quality, comparability and replicability of online research into key societal issues. This can also engender greater distrust between platforms and the wider research community.

Overall, there is a pressing need for a set of consistent baseline standards for researchers' access to platform data, both in terms of what data researchers are allowed to access and how it can be accessed; this is especially the case with research into new or emerging online platforms. Such standards would help to level the playing field and reduce the current asymmetries between platforms and researchers.

Policy Context

Outside of this project, we have also encountered similar challenges with conducting research on a range of other emerging platforms across the constantly evolving online ecosystem.[63] As the platform scoping exercise we conducted in Phase I of this project demonstrated, there are a wide range of other smaller- and medium-sized online platforms that are also being adopted by extremists and other harmful actors. We suspect this trend is only likely to increase if larger platforms become less hospitable for these kinds of online communities and activities, especially as regulation like the EU's Digital Services Act (DSA) comes into force.[64] It is also possible that, at least in the short-term and/or in jurisdictions without regulation requiring data access, some platforms may move to further restrict or roll back existing data access options. Without strong regulatory data access requirements, platforms may have an incentive to avoid the scrutiny that independent research can provide or attempt to profit from the provision of data access; for example, in February 2023, Twitter's announced changes restricting free access to the Twitter API, and it has been reported that Meta has plans to shut down existing access to data via the Meta-owned CrowdTangle.[65]

While many proposed pieces of regulation primarily focus on the largest and most dominant platforms, changes to the regulatory environment for online platforms may still impact smaller- and medium-sized online platforms, albeit to a lesser extent. Depending on the jurisdiction, these platforms may still be required to provide more comprehensive and consistent TOS, additional transparency and/or better access to data for regulators or third-party researchers. They may also need to have systems in place to quickly and effectively address reports of illegal content or activity on their platforms. In the case of Telegram, its considerable growth in recent years could soon make it subject to the same requirements as other more established large platforms, compliance with which would require significant changes to how it currently operates. In the coming years, if medium-sized platforms like Discord continue to grow, then they may also face additional regulatory obligations that require greater investments in content moderation, transparency and data access.

Smaller platforms like Odysee, which have positioned themselves as 'free speech' alternatives to larger, more mainstream tech companies, are also likely to face a choice; if they want to grow and move towards long-term profitability, they are likely to face higher regulatory requirements and scrutiny, as well as need additional technical and human resources to ensure their platform can handle increased activity. However, this may not be popular with their existing userbases, which often choose to adopt these platforms precisely because of their lax approaches to content moderation. Alternatively, if they seek to placate these types of users or communities by refusing to comply with regulatory or legal requirements, they may no longer be able to operate in certain jurisdictions and/or they may face legal action or financial penalties from regulators looking to reign in the levels of harmful content or behaviours they facilitate.

Examining the data access provisions and userbase thresholds in specific pieces of existing or proposed regulation in

key jurisdictions suggests that many of the challenges raised in this report will not be fully addressed. Much regulation relies on 'monthly active users' (MAUs) as a key metric to determine whether platforms meet the threshold for the strongest set of regulatory requirements; this means that many smaller platforms which nonetheless pose significant online risks would not be required to address existing barriers to data access. The following section provides an initial overview of existing or proposed pieces of regulation in key jurisdictions, and briefly assess whether they could have an impact on the barriers to research presented by smaller- and medium-sized online platforms. The final phase of this project will examine the policy implications of the data access challenges our research has identified in more detail.

EU: Digital Services Act (DSA) and Strengthened Code of Practice on Disinformation

Under the EU's DSA, passed in July 2022, platforms will be designated as 'very large online platforms' (or VLOPs) if they average 45 million or more MAUs across EU countries. These larger platforms will be required to provide real-time data access where technically possible, provided that the data is publicly accessible (Art. 40). This could include metrics like aggregated interactions with content from public pages, public groups or accounts, including impression and engagement data, such as the number of reactions, shares and comments. Access would need to be provided within a reasonable period upon request by the national-level regulator, the Digital Services Coordinator (DSC), in the EU country in which the company is established or the European Commission (EC). DSCs and the EC may only use the data provided for monitoring and assessing compliance with the DSA, and they must consider the rights and interests of providers and users, including the protection of personal data, confidential information (trade secrets) and maintaining the security of the service. Data would need to be provided through appropriate interfaces specified in the request from DCSs or the EC, including online databases or APIs.

Similarly, if requested by the DSC, data can be provided to vetted researchers to support the detection, identification and understanding of the systemic risks identified in the DSA in EU contexts. Vetted researchers must be affiliated with a research organisation, be independent from commercial interests and be able to meet certain data protection and security standards. Researchers will need to demonstrate the necessity and proportionality of their data access requests, disclose their funding and make any research freely and publicly available. Upcoming delegated acts (secondary EU legislation) will introduce further conditions, procedures and independent advisory mechanisms to support the sharing of data with external researchers. Platforms will need to anonymise or pseudonymise personal data unless doing so would make the intended (and legitimate) research objectives impossible.

In 2022, the EU also introduced a new Strengthened Code of Practice on Disinformation, a voluntary code with currently 34 signatories, including a range of organisations with an interest in combatting disinformation and large tech companies like Meta (Facebook and Instagram), Google (YouTube), Twitter, TikTok, Microsoft and Twitch, as well as a number of smaller companies, such as Clubhouse and Vimeo.[66] Under 'Section VI. Empowering the Research Community', the code contains a series of data access provisions, including Commitment 26 for signatory platforms to provide non-personal data and anonymised, aggregated or public data for research purposes on disinformation. As with the DSA, platforms should provide real-time (or near real-time) machine-readable data access through APIs or other open and accessible technical mechanisms as well as ensure reasonable tools and processes are in place to mitigate risks of abuse (such as, malicious or commercial uses of data). In order to qualify, proposed research must be compliant with ethical and methodological best practices, for example, those contained in the European Digital Media Observatory's (EDMO) draft Code of Conduct on Access to Platform Data, and research teams may include civil society as well as academic organisations.[67] The Strengthened Code of Practice on Disinformation also includes Commitment 27 to create an independent, third-party body to vet researchers and research proposals.

Of the platforms assessed in this report, none is likely to reach the threshold required in the DSA for these requirements to apply.viii As a result, while the DSA's requirements would go a long way to addressing current barriers to research on larger platforms, they may not help to solve many of the challenges outlined in this report related to smaller- or medium-sized platforms. Similarly, none of the platforms covered in this report are currently signed up to the voluntary Strengthened Code of Practice on Disinformation, meaning that it will also not address barriers to research on smaller and medium-sized platforms such as Telegram, Discord or Odysee unless or until they commit to its provisions.

viii At around 38.5 million, Telegram claims to have fewer users than the 45-million threshold for VLOPs in the EU. 'FAQ', Telegram, https://telegram.org/faq.

US: Platform Transparency and Accountability Act (PATA) and Digital Services Oversight and Safety Act (DSOSA)

Similarly to the DSA, proposed legislation in the US such as PATAix and DSOSAx both contain MAU thresholds to determine the obligations for different sized platforms. In PATA, platforms hosting user-generated content would need to have at least 25 million MAUs in the US to qualify (Sec. 2). The legislation would require the National Science Foundation (NSF) to identify the data and information that platforms must make available to qualified researchers. These have to be feasible for the platform to provide; proportionate to the needs of researchers to complete the research; and not create undue burdens for the platform (Sec. 4). Under PATA, qualified researchers must be university affiliated and submit applications to NSF for their specific research proposal (Sec. 2). NSF would also establish a process to solicit research applications from researchers and define guidelines and criteria to determine how to evaluate those applications (Sec. 4).

The bill would also create a Platform Accountability and Transparency Office (PATO) within the Federal Trade Commission (FTC) to notify platforms of research applications and establish privacy and cybersecurity safeguards for the use of the data in question, such as encryption or anonymisation of data to protect the privacy of individual users (Sec. 4). Researchers would be required to submit a pre-publication version of their research to the platform and PATO for evaluation to confirm that the analysis does not expose personal information, trade secrets or confidential commercial information (Sec. 5). PATA would also provide protection to researchers via the creation of a 'safe harbour' to prevent platforms from taking legal action against researchers who obtain information consensually and with other privacy protections in place (Sec. 11). Finally, the bill would also allow the FTC to require transparency reporting or disclosures to be made available in a form that is accessible and understandable to the public or accessible for analysis by researchers, journalists and the public via mechanisms, such as APIs. This could include ad libraries, information about widely disseminated content, information about content moderation decisions or information about algorithms (Sec. 12).

DSOSA, modelled to some extent on the EU's DSA, sets the threshold for 'large covered platforms' at 66 million MAUs and 10 million for 'covered platforms'. Similar to PATA, the bill would require the FTC to issue rules regarding the types of data that should be made available to certified researchers (Sec. 10 (c)), including the types of and mechanisms for data access, for example, considering the size and sampling techniques used to create the datasets. It would also require large covered platforms to provide a detailed ad library (Sec. 10(f)); access to key metrics for high-reach and high-engagement public content; and transparency on content that a platform amplifies (Sec. 10(g)). The FTC could choose to sponsor a Federally Funded Research and Development Center (or FFRDC) to facilitate information sharing between covered platforms and certified researchers. The FTC would also be required to ensure that data access does not infringe upon reasonable expectations of personal privacy of users (e.g. requiring platforms to anonymise any information that is not public content) and consider when privacy-preserving techniques, such as differential privacy and statistical noise, should be used (Sec. 10 (c)).

The bill would also create an Office of Independent Research Facilitation at the FTC that would certify researchers from academia and civil society to study the impact of content moderation processes, product design decisions and algorithms on society, politics, the spread of hate, harassment and extremism, security, privacy, and physical and mental health (Sec. 10 (a)). To qualify, research organisations would need to be an institution of higher education or a non-profit (501(c)(3)), which has a mission that includes developing a deeper understanding of the impacts of platforms on society; has the capacity both to comply with rules for secure researcher access and to use appropriate data science; and adopts investigative and qualitative research methods and best practices (Sec. 10 (b)). The bill would also introduce safe harbour provisions for certified researchers that create accounts solely for a research project or collect information provided for research purposes by a user (e.g. via a browser extension or plug-in) where the user has provided informed consent (Sec. 10(c)).

ix PATA – the text of the bill is available here. A section-by-section summary of the bill is available here.

x DSOSA – the text of the bill is available here. A section-by-section summary of the bill is available here.

As with the DSA in an EU context, the thresholds set in these proposals are likely to exclude many smaller-sized platforms. While the lower 10 million threshold for 'covered platforms' in DSOSA may capture some medium-sized platforms, such as Telegram or Discord, the most impactful data access requirements are aimed at large covered platforms. As a result, neither proposal would comprehensively address the barriers to research on smaller platforms encountered during our work for this report. It should also be noted that the likelihood of these federal bills being passed in Congress is low due to the greater executive interest in privacy and competition-focused legislation; there are also partisan disagreements over the path forward for social media-focused regulation (beyond regulation specifically aimed at protecting children's online safety, which may have stronger prospects). This suggests that the US may lag behind on regulation introducing new data access requirements in other jurisdictions, at least in the short to medium term.

UK: Online Safety Bill (OSB)

The proposed OSB in its current form (as of February 2023) still lacks clarity over a potential future data access regime. The Office of Communications (Ofcom), the designated regulator, would be able to request data from platforms and potentially share this with selected third-party organisations to conduct supportive research. Within two years of the legislation being adopted, Ofcom would also be required to produce a report describing how and to what extent those carrying out independent research into online safety are currently able to obtain information from regulated platforms. The report would also explore legal and other issues which currently constrain data access, as well as assess the extent to which greater access could be achieved. Following the publication of the report, Ofcom could choose to produce additional guidance for regulated platforms and researchers. In comparison to the DSA, the OSB contains relatively weak provisions and a lack of clarity on data access for researchers, which may not be clarified until several years into the new regulatory regime. It therefore remains unclear the extent to which, if at all, the OSB would address the barriers to data access raised in this report on large, medium or small platforms.

Germany: Network Enforcement Act (NetzDG)

NetzDG, first introduced in 2017 and then updated in 2021, is an example of an earlier generation of more content-centric digital regulation that is focused on the removal of illegal online content in Germany.[68] The regulation also has a userbase threshold for platforms to be in scope, which is set at two million or more registered users in Germany. The regulation includes various data access provisions (§ 5a) that allow researchers to request information on the dissemination of and engagement with content that is subject to complaints or removals, as well as access to the data used to train platforms' automated systems to detect illegal content. Any data requested must be necessary to conduct scientific research in the public interest (e.g. researching the nature, scope, causes and effects of public communication on social networks) but otherwise, there are no restrictions on the types of organisations or researchers that can make use of these provisions.

The regulation does, however, allow platforms to recoup reasonable costs associated with providing data per request, up to a limit of €5,000. Researchers must also provide a description of their approach to data protection, for example, on the steps taken to prevent data from being used for any other purposes or technical and organisational measures to ensure data security. Any data shared by platforms must be anonymised or pseudonymised wherever possible unless this prevents the purpose of the research. However, the precise mechanisms for sharing data (e.g. via an API) are not specified, and the data access provisions in NetzDG have not been widely used by researchers to date, potentially because of the associated costs. As a result, despite having been in place for several years, NetzDG has not proved to be effective in facilitating data access for researchers on large, medium or smaller platforms.

Recommendations

Based on the findings and implications for researchers outlined in this report, as well as the broader policy context set out above, we provide here a series of overarching recommendations for all online platforms, policy-makers and regulators, and civil society and academic researchers to address the range of technological, ethical and legal, and fragmentation barriers we have encountered during our research.

We acknowledge that some of our recommendations will create additional work for platforms, especially smaller platforms with fewer resources or technical capabilities. However, at present, platforms provide effective functionalities for enabling the communication and coordination of harmful actors while only offering restricted or poor functionalities for public interest research into these spaces. We argue this imbalance must be addressed.

Technological Barriers

Research into modern online spaces faces substantial challenges due to the speed and scale at which content can be created. Technological inability to access relevant data – whether the result of deliberate barriers, complicated platform structures or poorly documented technologies – substantially impacts public interest research, particularly for organisations with limited access to technical capabilities. The proliferation of novel technologies like blockchain will only increase this complexity and further limit the availability of technical tools and expertise.

- **Platforms should provide permissioned data access (e.g. via APIs) to third-party researchers conducting public interest research.** These should be accompanied by clear, consistent and accessible documentation that includes guidance on the types of data that can be collected and sufficient limits on the collection of sensitive personal data to protect user privacy. Where possible, platforms should also streamline the range of tools they directly provide to access data in order to reduce the burden of researchers needing to use multiple tools with overlapping capabilities.
- **Policy-makers and regulators should require data access that is accurate, reliable and sufficient for public interest research from platforms in future regulation. This would include the wide range of smaller- and medium-sized online platforms rather than the current focus which is simply on the largest.** This is vital to better understand how illegal or harmful online activity is increasingly conducted or coordinated on smaller- and medium-sized platforms that may present considerable risks despite their size. As far as possible, policy-makers should also work across jurisdictions to ensure consistency in data access requirements for platforms (which would help avoid over-burdening smaller businesses) and help develop best practices for both platforms and researchers over time.
- **Academic and civil society researchers should share effective data collection approaches or tools, as well as any lessons learnt, for accessing the growing range of platforms across the evolving online ecosystem.** On platforms that require more technical expertise to access data, considerable time can be spent exploring and testing potential options; the research community (and its funders) should seek to avoid duplicative or proprietary approaches. Wherever possible, they should instead pursue economies of scale and reduce disincentives for sharing methods, tools and approaches. This could be coordinated by cross-sectoral initiatives, such as the Coalition for Independent Tech Research, or via nascent independent data access bodies, such as those proposed by EDMO or the Carnegie Endowment for International Peace (CEIP).[69]

Fragmentation Barriers

Alongside the following specific recommendation related to systematic search functionalities on platforms, the recommendations for addressing technological barriers set out above would also help to address our identified fragmentation barriers.

- **Platforms should provide data access in a systematic way that enables researchers to reliably access accurate data from across public areas of the platform**, rather than having to manually explore the platform for spaces of communities of interest. This would allow for greater scale but also more targeted research as researchers would be able to identify relevant communities (and therefore data) more easily. Tools for systematic

search, whether platform-native or third-party, should demonstrate reliability, coverage and accuracy, which could be supported and assessed, for example, via third-party reviews conducted by researchers or potentially regulators.

Ethical Barriers

One of the key challenges raised by this research is how to protect the right to privacy while conducting public interest research on harmful communities and behaviours. Unfortunately, we lack a common definition of private spaces, making it difficult for researchers to know precisely which spaces to treat as private versus public. At the same time, the lack of a shared understanding of what constitutes a private space allows platforms to limit transparency and access to spaces that are easily accessible and arguably public.

As we have previously argued, factors such as size, purpose, accessibility and the nature of relationships between users of a channel or a community should be taken into account when making assessments about public or private spaces.[70] There is also a danger that illegal or harmful online activity openly conducted in nominally private online spaces could create uncertainty and undermine the case for the privacy (e.g. via encryption) of genuinely private online spaces and means of communications.

- **Platforms should determine a reasonable limit for the number of members that can participate in private groups and channels and declare online spaces with audiences over a certain threshold as public.** This should also help to provide greater clarity to both platform users and third-party researchers around what is acceptable in these spaces, as well as which types of data may be accessible to third parties with appropriate privacy and data protection safeguards. Content with no reasonable privacy expectations (e.g. content posted on public pages) should be made available via vetted API access, including relevant metrics on reach, impressions and engagement.
- **Policy-makers should consider introducing requirements for platforms to clarify which areas are truly public or private and set reasonable thresholds for the number of users that can participate in private online spaces if platforms are unwilling or disincentivised to do so voluntarily.** Rather than introducing blanket regulatory thresholds in this area, regulation could require companies to set clear and reasonable limits based on the nature of their platforms and risk assessments that consider a platforms' specific features or functionalities (e.g. encryption), risks and potential vulnerabilities. Regulation could also encourage companies to make clear to users which aspects of their platform are more public or more private, as well as the consequences of this for user privacy and researcher or third-party data access. An independent regulator could then assess, based on the risk assessment, whether the limit set by the platform is appropriate and sufficiently mitigates any risks or harms identified.
- **The research community should work to formalise ethical approaches to researching public, semi-private and private online spaces. This would be in line with the potential severity of the risks such spaces could pose.** Such approaches must balance rights to privacy with the rights of those that may be negatively impacted if these spaces, especially the largest and least moderated ones, remain out of the reach of researchers. Such efforts should build on the existing and well-established field of internet research ethics. Coordination could occur through existing initiatives, such as the Association of Internet Researchers (AoIR), or the potential future independent data access bodies mentioned above (e.g. EDMO).[71]

Legal Barriers

- **Platforms should provide clear TOS, not only in regard to the types of content and activities they allow, but also to how they apply to researchers accessing data. These should then be enforced consistently.** TOS should allow for privacy-respecting public interest research rather than effectively insulating the platform from public scrutiny. Greater clarity in TOS would also set clear expectations regarding the use of data collected by third-parties; this would provide more certainty for researchers as well as help users to understand how their data may be used.
- **Policy-makers should introduce legal protections for researchers conducting public interest, privacy-respecting online research.** Where researchers use proportionate methods and approaches and make sufficient

efforts to protect user privacy, they should not face the potential risk of legal action from platforms when researching important societal questions around the extent and impact of illegal or harmful activity in public spaces online. This is particularly relevant where platform TOS are incomplete, unclear or ambiguous. Platforms should not be able to completely evade scrutiny by placing a blanket ban on data access in their TOS. In-lieu of governments introducing legal protections for researchers, responsible platforms should establish voluntary exemptions in their TOS to permit research via methods like crowdsourced data collection (where researchers have secured the informed consent of users).

- **Academic and civil society researchers should also consider opportunities to share or pool expertise on the legal implications of platform data access.** The legal review of a platform's TOS and other related conditions or policies can be time consuming, resource intensive and often jurisdictionally specific. Given that many researchers may not have access to outside legal support (spanning data protection and contract law), greater coordination and sharing of resources and expertise could help to reduce legal risks and barriers to entry for online research, thereby increasing equity among researchers. Again, as with the technological and ethical recommendations above, this type of cross-sectoral cooperation should be facilitated by existing initiatives where possible.

Looking Ahead

Overall, the current prospects for researchers' access to data from platforms across the evolving online ecosystem are mixed. While existing and incoming regulation in key jurisdictions offers considerable promise in terms of increasing platform transparency and broadening external data access, particularly for the largest platforms, as we have noted, there are also incentives for platforms not to provide or to further restrict data access in many contexts. It remains to be seen whether regulations in key jurisdictions like the EU will have spill-over benefits and whether platforms will decide to extend similar access in contexts where they are not legally required to.

Much of this regulation will also not substantially improve the status quo regarding many smaller- or medium-sized platforms which will not be subject to the more extensive data access requirements. This will likely result in a continuation of the disempowerment researchers face relative to many platforms, with few guarantees that this situation will not deteriorate. For example, some companies may attempt to restrict data access via technological or legal means in order to avoid further regulatory or independent scrutiny into the risks and harmful content or behaviours present on their platforms. Regardless of whether the prospects for data access improve, the online ecosystem will inevitably continue to evolve at pace, increasing the number, diversity and complexity of online platforms that researchers may wish to access.

To progress beyond the current situation, at a minimum, policy-makers should consider introducing legal exemptions or protections for researchers conducting public interest and privacy-respecting online research. Policy-makers could also consider how to better incentivise and support platforms through non-regulatory means to voluntarily provide tools that offer structured and accurate data, especially for platforms that do not meet regulatory thresholds that require greater data access. In a more ambitious scenario that significantly reduces barriers to research and may require further regulation in many contexts, platforms would provide permissioned, standardised and privacy-preserving data access to researchers, with guaranteed minimum standards of accuracy and clear TOS that permit public interest research. Smaller platforms may need to be supported with resources to facilitate this, alongside the development of a recognised intermediary body that can ensure the ongoing quality of data and make it available in a standardised and privacy-preserving fashion to researchers. Researchers, as well as the funders of online research, should also consider how to create greater economies of scale across the field; this would reduce the current barriers to entry that slow progress towards a more equitable, diverse and global research community.

Endnotes

1 Guhl, Jakob, Marsh, Oliver and Tuck, Henry, 'Researching the Evolving Online Ecosystem: Barriers, Methods and Future Changes', *Institute for Strategic Dialogue*, July 2022, https://www.isdglobal.org/isd-publications/researching-the-evolving-online-ecosystem-barriers-methods-and-future-challenges/.

2 'Extracts From ISD's Submitted Response to the UK Government Online Harms White Paper', *Institute for Strategic Dialogue*, July 2019, https://www.isdglobal.org/isd-publications/extracts-from-isds-submitted-response-to-the-uk-government-online-harms-white-paper/.

3 Channels FAQ, *Telegram*, https://telegram.org/faq_channels

4 '700 Million Users and Telegram Premium', *Telegram*, June 2022, https://telegram.org/blog/700-million-and-premium.

5 'Terms of Service', *Telegram* https://telegram.org/tos.

6 'Europol and Telegram take terrorist propaganda online', *Europol*, November 2019, https://www.europol.europa.eu/media-press/newsroom/news/europol-and-telegram-take-terrorist-propaganda-online.

7 Gerster, Lea et al, 'Stützpfeiler Telegram. Wie Rechtsextreme und Verschwörungsideolog:innen auf Telegram ihre Infrastruktur ausbauen', *Institute for Strategic Dialogue*, December 2021, https://www.isdglobal.org/isd-publications/stutzpfeiler-telegram-wie-rechtsextreme-und-verschworungsideologinnen-auf-telegram-ihre-infrastruktur-ausbauen/.

8 'FAQ – Q: What's the difference between groups and channels?', *Telegram*, https://telegram.org/faq#q-what-39s-the-difference-between-groups-and-channels.

9 'Ratings of Telegram groups', *TGStat*, https://tgstat.com/ratings/chats.

10 'Belarus Carries Out Wave Of Detentions For Subscribing to "Extremist" Telegram Channels', *Radio Free Europe/Radio Liberty*, October 2021, https://www.rferl.org/a/belarus-telegram-extremist-detentions/31530720.html.

11 'Через сервис TGStat стали доступны данные о протестных приватных чатах и их участниках', *Mediazona*, April 2022, https://mediazona.by/news/2022/04/28/tgstat.

12 Frühling, Milla, 'OLIVER JANICH – QANON-DESINFORMATIONEN AUF ALLEN KANÄLEN', *Belltower.News*, October 2020, https://www.belltower.news/social-media-rechtsaussen-oliver-janich-qanon-desinformationen-auf-allen-kanaelen-105577/.

13 'TELEGRAM-KANAL', *Oliver Janich Investigativ*, https://www.oliverjanich.de/telegram-messenger.

14 Institute for Strategic Dialogue, July 2019, op. cit.

15 Gerster et al, December 2022, op. cit.

16 Rogers, Iain, 'Germany Foils Plot to Sabotage Democracy, Kidnap Health Minister', *Bloomberg UK*, April 2022, https://www.bloomberg.com/news/articles/2022-04-14/germany-foils-plot-to-sabotage-democracy-kidnap-health-minister.

17 'Raid due to murder plans against Kretschmer', ZDFheute, December 2021, https://www.zdf.de/nachrichten/politik/razzia-telegram-mordplan-kretschmer-100.html.

18 'Täglich Tötungsaufrufe auf Telegram', Tagesschau, January 2021, https://www.tagesschau.de/investigativ/funk/ todesdrohungen-telegram-101.html.

19 'Ermittler decken illegalen Handel auf', Tagesschau, October 2020, https://www.tagesschau.de/inland/telegram-illegaler-handel-105.html.

20 See for example: Sold, Manjana and Junk, Julian, 'Researching Extremist Content on Social Media Platforms: Data Protection and Research Ethics Challenges and Opportunities', *Global Network on Extremism and Technology*, 2021, https://gnet-research.org/wp-content/uploads/2021/01/GNET-Report-Researching-Extremist-Content-Social-Media-Ethics.pdf; Conway, Maura, 'Online Extremism and Terrorism Research Ethics: Researcher Safety, Informed Consent, and the Need for Tailored Guidelines, Terrorism and Political Violence', *Terrorism and Political Violence*, 33 (2), pp. 367–380, March 2021, 10.1080/09546553.2021.1880235; "Internet Research: Ethical Guidelines 3.0', AoIR, October 2019, https://aoir.org/reports /ethics3.pdf; 'A Guide to Internet Research Ethics', NESH, June 2019 https://www.forskningsetikk.no/en/guidelines/social-sciences-humanities-law-and-theology/a-guide-to-internet-research-ethics/.

21 'PROJEKT RADIKALISIERUNG IN RECHTSEXTREMEN ONLINE-SUBKULTUREN ENTGEGENTRETEN', *Institute for Strategic Dialogue Germany*, https://isdgermany.org/projekt-bmj/.

22 'Nazi-Song ist Antifa-Experiment', Taz, January 2022, https://taz.de/Rechte-Musik-auf-Streaming-Plattformen/!5828978/.

23 Dockery, Wesley, 'Germany cracks down on far-right Telegram users', *Deutsche Welle*, September 2022, https://www.dw.com/en/germany-cracks-down-on-far-right-telegram-users/a-60715438.

24 'FAQ – Q: There's illegal content on Telegram. How do I take it down?', *Telegram*, https://telegram.org/faq#q-there-39s-illegal-content-on-telegram-how-do-i-take-it-down.

25 'Antwort der Landesregierung auf eine Kleine Anfrage zur schriftlichen Beantwortung', *Landtag Von Sachsen-Anhalt*, https://www.landtag.sachsen-anhalt.de/fileadmin/files/drs/wp8/drs/d1519aak.pdf.

26 Efforts to delegitimise the state, Bundesamt für Verfassungsschutz, https://www.verfassungsschutz.de/EN/topics/efforts-to-delegitimise-the-state/efforts-to-delegitimise-the-state_node.html

27 'Mit Haftbefehl gesuchter Hetzer Attila Hildmann sieht sich als Nachfolger von Adolf Hitler', RTL, October 2022, https://www.rtl.de/cms/attila-hildmann-von-stern-reportern-in-der-tuerkei-aufgespuert-er-wird-mit-haftbefehl-gesucht-5013127.html.

28 Steinke, Ronen, 'Office for the Protection of the Constitution: Alone among false friends', *Süddeutsche Zeitung*, September 2022, https://www.sueddeutsche.de/projekte/artikel/politik/verfassungsschutz-rechtsextreme-social-media-telegram-virtuelle-agenten-reichsbuerger-coronaleugner-rassismus-antisemitismus-verschwoerungsideologie-e222942/?reduced=true.

29 Curry, David, 'Discord Revenue and Usage Statistics (2023)', *Business of Apps*, January 2023, https://www.businessofapps.com/data/discord-statistics/.

30 'Top 100 Biggest Discord Servers', *Discord*, https://discords.com/servers/top-100.

31 'Discord servers tagged with 4Chan', *Discord*, https://disboard.org/servers/tag/4chan.

32 Roose, Kevin, 'This Was the Alt-Right's Favorite Chat App. Then Came Charlottesville', *The New York Times*, August 2017, https://www.nytimes.com/2017/08/15/technology/discord-chat-app-alt-right.html; Davey, Jacob, and Ebner, Julia, 'The Fringe Insurgency – Connectivity, Convergence and Mainstreaming of the Extreme Right', I*nstitute for Strategic Dialogue*, October 2017, https://www.isdglobal.org/isd-publications/the-fringe-insurgency-connectivity-convergence-and-mainstreaming-of-the-extreme-right/.

33 Alexander, Julia, 'Discord is purging alt-right, white nationalist and hateful servers', *Polygon*, February 2018, https://www.polygon.com/2018/2/28/17061774/discord-alt-right-atomwaffen-ban-centipede-central-nordic-resistance-movement.

34 Allyn, Bobby, 'Group-Chat App Discord Says It Banned More Than 2,000 Extremist Communities', *NPR*, April 2021, https://www.npr.org/2021/04/05/983855753/group-chat-app-discord-says-it-banned-more-than-2-000-extremist-communities.

35 Gallagher, Aoife et al, 'Gaming and Extremism: The Extreme Right on Discord', *Institute for Strategic Dialogue*, August 2021, https://www.isdglobal.org/isd-publications/gaming-and-extremism-the-extreme-right-on-discord/.

36 Ibid.

37 Ayad, Moustafa, 'Islamogram: Salafism and Alt-Right Online Subcultures', *Institute for Strategic Dialogue*, November 2021, https://www.isdglobal.org/isd-publications/islamogram-salafism-and-alt-right-online-subcultures/.

38 Wong, Wilson, 'Discord chat app faces moderation questions after mass shooting suspects are linked to platform', *NBC News*, July 2022, https://www.nbcnews.com/tech/tech-news/highland-park-shooting-suspect-bobby-e-crimo-iii-discord-server-raises-rcna36659.

39 Joyce, Kathryn and Lorber, Ben, '"Traditional" Catholics and white nationalist "groypers" forge a new far-right youth movement', *Salon*, May 2022, https://www.salon.com/2022/05/13/trad-catholics-and-nationalist-groypers-forge-a-new-far-right-youth-movement/.

40 Disboard, https://disboard.org/.

41 'Channel Resources', *Discord*, https://discord.com/developers/docs/resources/channel#get-channel-messages; "Dracovian/Discord-Scraper', Github (via Internet Archive), https://web.archive.org/web/20220727182403/https://github.com/Dracovian/Discord-Scraper.

42 'Terms of Service', *Discord*, https://discord.com/terms.

43 'Discord Developer Policy', *Discord*, https://discord.com/developers/docs/policies-and-agreements/developer-policy. Note: Discord's Developer Policy was updated in September 2022, and now includes the following in place of the previous wording referenced in the text: "You may not mine or scrape any data, content, or information available on or through Discord services".

44 Schlegel, Linda, 'Jumanji Extremism? How games and gamification could facilitate radicalization processes,' *Journal for Deradicalization*, 23, Summer 2020, pp. 1–44.

45 Vogel, William, 'People-first video platform Odysee Launches out of Beta, Enabling Creators to Reclaim Power and Monetization', *Cision PRWeb*, April 2023, https://www.prweb.com/releases/peoplefirstvideoplatformodyseelaunchesoutof-betaenablingcreatorstoreclaim-powerandmonetization/prweb17586549.htm.

46 Odysee (@OdyseeTeam), 'Big tech censorship has gone too far! Use Odysee as your alternative to YouTube. We're not perfect, but getting better. FYI. Livestream with us, and make more money than yt/twitch in donations,' *Twitter*, 30 September 2021, https://twitter.com/OdyseeTeam/status/1443642146280509464.

47 Odysee (@Odysee), 'Odysee & Freedom of the Press: A Letter From Our CEO', *Odysee*, March 2022, https://odysee.com/@Odysee:8/freedomofthepress:0.

48 Matlach, Paula, Hammer, Dominik and Schwieter, Christian, 'Auf Odysee: Die Rolle von Blockchain-Technologie für die Monetarisierung im rechtsextremen Onlinemilieu', *Institute for Strategic Dialogue*, August 2022, https://www.isdglobal.org/wp-content/uploads/2022/08/ISD_auf-odyssee_220810_digital.pdf.

49 Odysee (@OdyseeTeam), 'We aren't alt tech, we're new tech Alt tech = inferior version of what's come before with no selling points other than free speech New tech = better version of what's come before with the added benefit of free speech,' *Twitter*, 3 August 2021, https://twitter.com/OdyseeTeam/status/1422629329905848322.

50 LBRY, https://lbry.com/.

51 'Jeremy Kauffman: Leadership For New Hampshire', https://jeremy4nh.com/home/

52 Matlach, Paula, Hammer, Dominik and Schwieter, Christian, August 2022, op. cit.

53 'New breed of video sites thrive on misinformation and hate', *Rappler*, August 2022, https://www.rappler.com/technology/features/new-breed-video-sites-thrive-misinformation-hate/.

54 Marshall, Andrew R. C. and Tanfani, Joseph, 'New breed of video sites thrives on misinformation and hate', *Reuters*, August 2022, https://www.reuters.com/investigates/special-report/usa-media-misinformation/.

55 'Sign-up', *Odysee*, https://odysee.com/$/signup?redirect=/.

56 Odysee (@Odysee), 'Creators Will Be Earning From Ads Soon', *Odysee*, October 2021, https://odysee.com/@Odysee:8/creator-earnings:7.

57 Talley, Ian, 'Islamic State Turns to NFTs to Spread Terror Message', *The Wall Street Journal*, September 2022, https://www.wsj.com/articles/islamic-state-turns-to-nfts-to-spread-terror-message-11662292800.

58 'Terms of Service', *Odysee*, https://odysee.com/$/tos.

59 'Les origins de l'extrême-droite', *herodote.net*, October 2022, https://www.herodote.net/Les_origines_de_l_extreme_droite-article-2715.php.

60 'Odysee.com', Similar Web (via Internet Archive), https://web.archive.org/web/20221018123152/https://www.similarweb.com/website/odysee.com/; 'Youtube.com', Similar Web (via Internet Archive), https://web.archive.org/web/20221018122916/https://www.similarweb.com/website/youtube.com/#overview.

61 'Declaration of Indifference: Community Guidelines', *Odysee*, February 2022, https://help.odysee.tv/communityguidelines/.

62 See for example: Mir, Rory and Doctorow, Cory 'Facebook's Attack on Research is Everyone's Problem', *EFF*, August 2021, https://www.eff.org/deeplinks/2021/08/facebooks-attack-research-everyones-problem.

63 See for example: Hammer, Dominik, Gerster, Lea and Schwieter, Christian, 'Inside the Digital Labyrinth: Right-Wing Extremist Strategies of Decentralisation on the Internet & Possible Countermeasures', *Institute for Strategic Dialogue*, February 2023, https://www.isdglobal.org/wp-content/uploads/2023/02/Inside-the-Digital-Labyrinth.pdf.

64 'Digital Services Act', *European Parliament*, https://www.europarl.europa.eu/doceo/document/TA-9-2022-0269_EN.pdf.

65 'Twitter API Changes Set to Disrupt Public Interest Research', *Tech Policy Press*, February 2023, https://techpolicy.press/twitter-api-changes-set-to-disrupt-public-interest-research/; Lawler, Richard, 'Meta reportedly plans to shut down CrowdTangle, its tool that tracks popular social media posts', *The Verge*, June 2023, https://www.theverge.com/2022/6/23/23180357/meta-crowdtangle-shut-down-facebook-misinformation-viral-news-tracker.

66 'Signatories of the 2022 Strengthened Code of Practice on Disinformation', *European Commission*, June 2022, https://digital-strategy.ec.europa.eu/en/library/signatories-2022-strengthened-code-practice-disinformation.

67 'EDMO releases report on researcher access to platform data', *European Digital Media Observatory*, May 2022, https://edmo.eu/2022/05/31/edmo-releases-report-on-researcher-access-to-platform-data/.

68 'Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act, NetzDG) – Basic Information (2017)', *Bundesministerium der Justiz*, https://www.bmj.de/DE/Themen/FokusThemen/NetzDG/NetzDG_EN_node.html.

69 *Coalition for Independent Technology Research*, https://independenttechresearch.org/; European Digital Media Observatory, May 2022, op.cit.; Wanless, Alicia and Shapiro, Jacob N., 'A CERN Model for Studying the Information Environment', *Carnegie Endowment for International Peace*, November 2022, https://carnegieendowment.org/2022/11/17/cern-model-for-studying-information-environment-pub-88408.

70 Institute for Strategic Dialogue, July 2019, op. cit.

71 'Ethics', *AoIR*, https://aoir.org/ethics/.

ISD | Powering solutions to extremism, hate and disinformation

Amman | Berlin | London | Paris | Washington DC