Evidence Gathering for Platform Accountability Project Report

September 2023

Digital Freedom Fund

ABOUT DFF

The Digital Freedom Fund (DFF) exists to support the digital rights community in Europe to advance and protect human rights in digital spaces and reduce the negative impact of technology in the world. Since 2017, we have built a strong record of conscientious grant making, curated events, and an expanding network of partners and allies, all united in the desire to create systemic change.

OUR TEAM

This report was compiled with the help and contributions of many participants and members of DFF. Four main contributors are recognised below.

ALAYNA KENNEDY

The Technology Community Lead and project lead for this report, Alayna has experience as a data scientist and machine learning developer and has worked on responsible technology and AI ethics for years.

NIKITA KEKANA

The Senior Legal Officer at DFF, Nikita has extensive experience in digital and media law, international law and human rights. She has worked within the strategic legal defence team at a leading international non-governmental organisation and two of South Africa's big 5 law firms. Nikita holds an LLM cum laude in Public International Law from Leiden University.

CÉSAR MANSO-SAYAO

The Legal Officer at DFF, César is a law graduate from the University of Costa Rica, holds two master's degrees from the University of Barcelona, and has a postgraduate degree in Technopolitics and Rights in the Digital Age from Pompeu Fabra University.

MAURICIO LAZALA

The Director of the Digital Freedom Fund. He is an international human rights lawyer and advocate with extensive experience in leading on strategic, regional and programmatic areas of work.

THANKS TO PARTICIPATING ORGANISATIONS

We extend our gratitude to the organisations that actively participated in our year-long study on evidence gathering for platform accountability. From our three in-person events to our online forums, round tables, and feedback sessions, each of these groups helped contribute to the findings of this report. Their dedication, insights, and contributions have played a pivotal role in shaping our understanding of this issue. As an organisation deeply committed to a participatory approach and collective learning, DFF is proud to have collaborated with each of these groups and thankful for their invaluable support and collaboration.

AI Forensics	HestiaLabs
The Alan Turing Institute	Irish Council for Civil Liberties
AlgorithmWatch	Liberties
Amnesty Tech	Mnemonic
Article 19	The Mozilla Foundation
BEUC	noyb
The Binary	nr16
Bits of Freedom	Panoptykon Foundation
Connected by Data	PersonalData.io
Consumer Reports	Reset Tech
ConsciousDigital.org	Science & Design
CORRECTIV	Simply Secure
Data Rights	Stiftung Neue
DATA WO	Verantwortung e.V
Electronic Frontier	Tracking Exposed
Foundation (EFF)	University of Edinburgh
Ethical Intelligence	University of Groningen
Gesellschaft für Freiheitsrechte e.V. (GFF)	

CONTENT

EXECUTIVE SUMMARY

This report conveys the results of a year-long investigation into the ways that evidence gathering can assist in platform accountability advocacy and litigation.

Large technology companies lack transparency, and to successfully hold them accountable for digital rights harms, we need better methods of gathering information about their inner workings, algorithms, and data processes. As these platforms wield unprecedented power, it is imperative to bridge the information gap through evidence-driven insights to ensure a balanced, transparent, and accountable digital environment.

After three in-person events and many online roundtables, feedback sessions, and interviews, we found that there is a shared vision for platform accountability work among litigators and technologists, and that opportunities exist to further pursue platform accountability through litigation and cooperation of these groups. We also found that many tools and projects for evidence-based litigation already exist, and that a stronger community is needed to disseminate and maintain these existing resources.

Looking to the future, our recommendations for the field include building knowledge-sharing mechanisms like a collaborative knowledge hub and a community of practice between litigators and technologists. Further solutions involve technical projects like a data sharing pipeline that litigators can use to share information and evidence for their cases. In addition to sharing data, organisations can learn from technologists by developing better data collection strategies and educating themselves about the basics of machine learning, natural language processing, and data collection techniques like Application Programming Interface (API) scrapes. **Finally, we also recommend investing in non-law remediation efforts and user-centric mechanisms to prevent and combat the harms of large technology platforms.**

INTRODUCTION

What are digital rights?

As society has increasingly become datafied, new frameworks of activism have emerged for dealing with the harms enacted by a technological society. While activists have fought for justice and access to human rights for many years, new forms of discrimination have been introduced with recent data processing and algorithmic technologies. **A new framing of justice, data justice, has evolved to establish fairness and justice in the ways people are represented and treated as a result of their digital data.[1]** For the last few decades, digital rights have evolved, expanded, and been challenged by the rapid advancement of technology and the ways that the digital world has infringed on all our lives, creating a need for a new field of human rights work: digital rights.[2] The digital rights field emerged to ensure that human rights would be upheld online and in technologically mediated environments, and digital rights now broadly addresses concerns about internet companies' use of personal data, surveillance technology, and algorithmic bias and discrimination, among others.[3]

> At DFF, we believe that digital rights are human rights, which means that most if not all human rights have a digital dimension to them. Our 2020 report, Digital Rights are Human Rights, explores this connection in more detail.

What is platform accountability?

While early efforts in the digital rights field during the late 1990s and early 2000s focused on ensuring basic online freedoms, such as freedom of expression and access to information, the emergence of social media platforms and a data-based economy has shifted much of the **digital rights focus towards the harms produced by large technology companies.** In an economy overwhelmingly characterised by data-for-profit, social media and search platforms like Facebook, Google, and Instagram have become some of the world's most profitable companies through their collection and sale of user data. [4] Other platforms, like Uber & Tinder, built their business model on data-driven algorithms that promised to transform their industries. [5]

Today, these large technology platforms are everywhere. TikTok feeds us entertainment, Airbnb helps us plan vacations, Amazon serves as our shopping centre, and Doordash, Deliveroo, and Wolt literally feed us with food delivery. When Elon Musk bought Twitter (renamed X), he explicitly began to expand the platform to be an all-encompassing communications platform, trying to position it as the **"town square"** of information, and platforms like WhatsApp, Facebook, and Instagram are used by billions to share aspects of
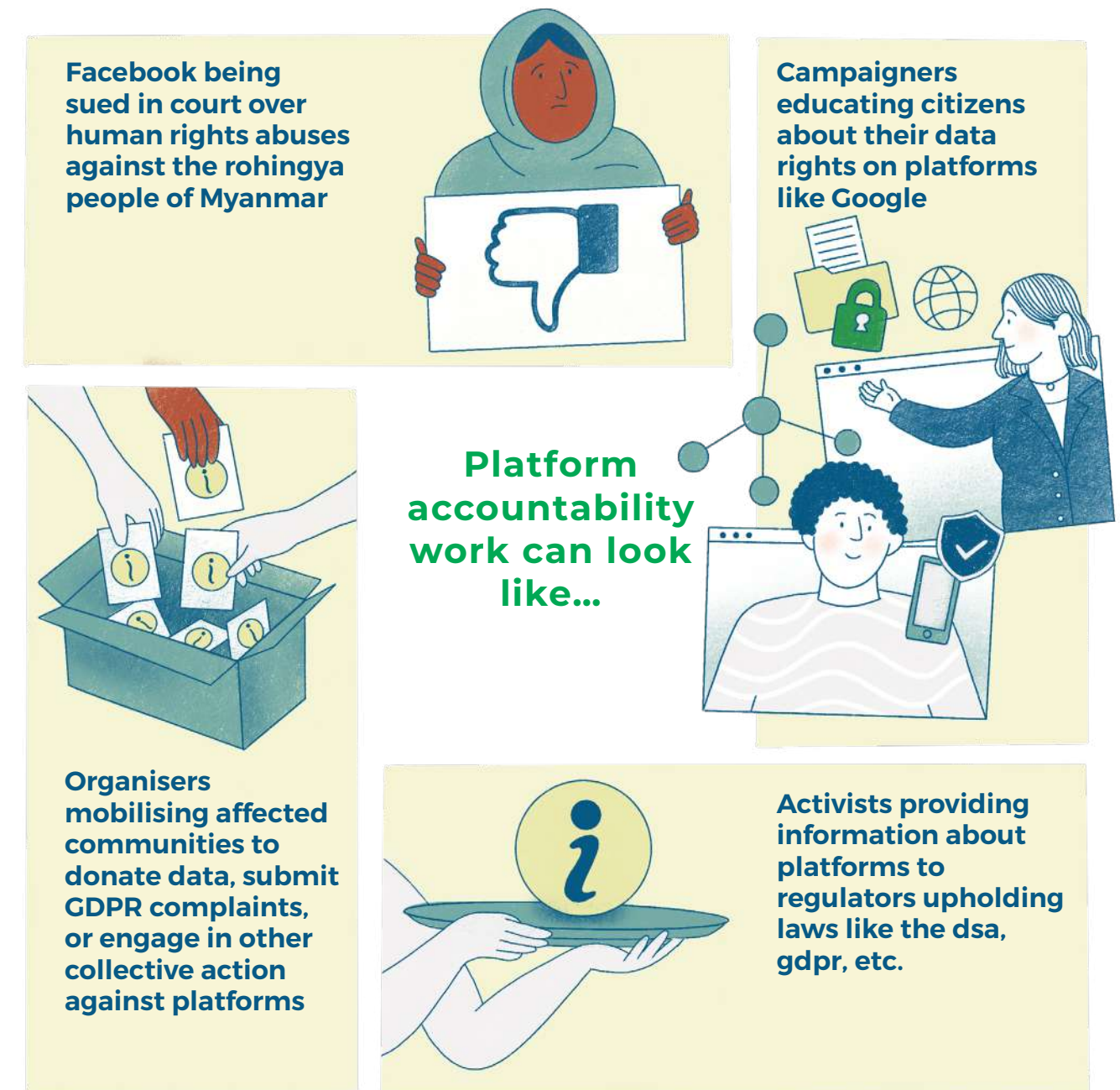
our personal lives. These platforms grew very rapidly without conducting prior human rights impact assessments or social risk analysis of their products. At the same time, government regulation has been unwilling, unable or very slow in creating accountability frameworks. our personal lives.

Much of the time, these platforms use artificial intelligence (AI) and machine-learning (ML) tools to form their recommendations and services.[6]

Rarely do platforms that use AI and ML provide transparency about how user data and AI interact to shape the content on the platform. Platforms use non-transparent technology practices in several broad areas: to moderate the user-generated content we see, to show us search results, news, and other information, to target and deliver advertisements, and to shape and curate all our user experiences on online platforms.[7]

Unsurprisingly, the prevalence and the secrecy of modern platforms has resulted in a suite of harms and digital rights violations. Facebook's ad delivery system discriminated against users based on their race, colour, religion, sex, disability, familial status and national origin.[8] Whistleblowers like Frances Haugen exposed the ways that social media companies actively support measures that negatively affect their users to generate more advertisement revenue.[9] Investigations into gig-work platforms like Uber & TaskRabbit have revealed that they often pay their gig workers below the minimum wage and provide insecure income.[10, 11] Furthermore, social media platforms have contributed to misinformation, propagating hate speech, and exacerbating political tensions during the last decade.[12]

To hold platforms accountable means to ensure that they are abiding by local and regional laws, like the European Union (EU)'s General Data Protection Regulation (GDPR) and Digital Services Act (DSA). Mechanisms of platform accountability range from international human rights treaties and frameworks to soft law instruments like the UN Guiding Principles on Business and Human Rights and ethical frameworks such as the Organisation for Economic Cooperation and Development (OECD) AI Principles or the High-Level Expert Group on Artificial Intelligence Ethics Guidelines for Trustworthy AI. While these formal frameworks are important channels through which platform accountability occurs, at its core platform accountability is a broader effort to uphold users' digital rights, safety, and that when harms occur, they are addressed, reparations are offered, and harms are remediated.[13]

Facebook being sued in court over human rights abuses against the rohingya people of Myanmar

Campaigners educating citizens about their data rights on platforms like Google

Platform accountability work can look like...

Organisers mobilising affected communities to donate data, submit GDPR complaints, or engage in other collective action against platforms

Activists providing information about platforms to regulators upholding laws like the dsa, gdpr, etc.

Platforms growing power makes transparency important

Harms enacted by online platform are exacerbated by the lack of transparency into these platforms' use of user data and algorithmic systems.[14-16] As these platforms have expanded their influence over information dissemination, public discourse, and economic life, they have begun to assume many of the roles previously held by government actors. The unprecedented power of these platforms means they have the ability to shape public opinions, impact elections, and dictate business practices.[17]

In the absence of robust regulatory frameworks, their unchecked power has given rise to a multitude of challenges, including digital rights infringements, algorithmic biases, and the propagation of misinformation.

The urgency for transparency within these powerful platforms is amplified by the scale of their impact.

Knowing how these platforms work is an important step in ensuring that they uphold digital rights, fairness, and democratic processes. When platforms wield comparable authority to some country governments, the lack of transparency into their operations raises concerns about their internal decision-making processes.[14] To achieve effective platform accountability, evidence about their inner workings is needed. Robust evidence about platforms' policies, practices, and algorithms would serve many purposes, including being the cornerstone for informed strategic litigation, advocacy efforts, and policymaking. By gaining insight into their operations and data practices stakeholders can effectively challenge digital rights violations and demand accountability.

Here it is important to acknowledge that while the lack of transparency represents an important element of platform accountability, there are many other aspects around platforms' practices which are relevant for platform accountability. Some of these other aspects include platforms' failure to invest enough resources and personnel to deal with culture-specific nuances and content in Global Majority areas which drives the spread of hate speech and censorship; anti-competitive practices and market dominance dynamics exacerbating concentration of wealth and social inequality; excessive data collection without clear and informed consent propagates deceptive and manipulative dark patterns; and the detrimental environmental impacts of the data centres and energy-intensive data processing operations platforms rely on. All of these elements of platform accountability work go beyond the simple lack of transparency into these systems. While much excellent work is being done on these other elements, this report's goal is not to give a comprehensive picture of the work being done around platform accountability, but to provide insights on improving evidence gathering to do better platform accountability work, of all kinds. **Since lack of transparency lies at the heart of many of the problems enacted by platforms, it will be the central focus of this report.**

Need for evidence

As an actor in the digital rights field, DFF funds and supports litigation and accompanying campaigning and activism efforts on platform accountability within the Council of Europe. However, platform accountability work faces an enormous knowledge gap – because of the lack of transparency into the inner workings of online platforms. Litigation against these companies and legal recourse against the harms they perpetuate becomes extremely difficult, as we heard from many of our participants in this project. Litigators, campaigners, and other activists need more evidence to hold platforms accountable, and that evidence can be hard to find when platforms are reluctant to reveal any information about the data they hold or their algorithms.

In April of 2022, DFF attended a summit in collaboration with Reset and Aspiration to discuss the challenge of gathering evidence in support of platform accountability work. The objective was to explore potential collaborations that could strengthen platform accountability efforts, specifically addressing the issue of how to gather evidence about platforms when those platforms lack transparency. **Participants in the summit voiced frustrations about the scarcity of evidence and the power disparity between advocacy groups and technology companies.** This power imbalance makes it difficult to access the necessary evidence to advance accountability initiatives. The summit provided a platform for these issues to be discussed and potential paths forward to be explored.

CHALLENGES

· Non-transparent company policies.

· Scarcity of technical expertise and resources in digital rights community.

· Need for more involvement with technical research & investigation.
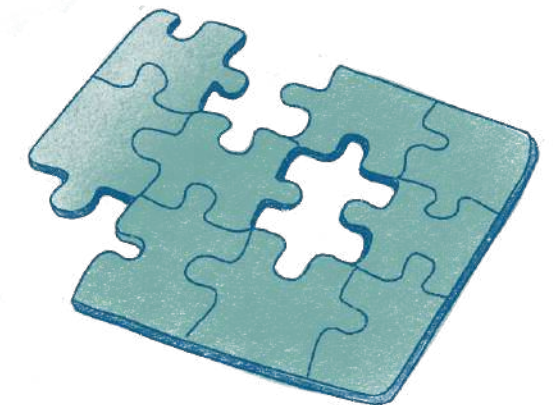
TAKEAWAYS

· Need within the digital rights community for more technical involvement.

· Gathering technical evidence is a key part of platform accountability.

· Digital rights could benefit from a stronger link between technical and litigation work.

· Non-transparent company policies.

· Scarcity of technical expertise and resources in digital rights community.

· Need for more involvement with technical research & investigation.

This summit began the process of envisioning and mapping out a framework for collaborative research and investigation between litigators and technologists. In addition, it laid the groundwork for a better understanding amongst the stakeholders who would use such a framework (researchers, lawyers, campaigners, etc.) and what their respective needs and challenges are in platform accountability work. The foundations for our work this past year, and this report, were laid out at that initial Luminate and Aspiration Summit in April 2022.

Objectives of this project

Following the summit, DFF began a year-long investigation into the issue of evidence gathering in platform accountability work, with the support of Luminate, Mozilla, Reset, and our digital rights community. We aimed to bring together litigators and activists with technologists to gather more and better evidence about the inner workings of platforms and the effects of their prevalence, practices and business models on users and society at large, **with the ultimate goal of holding these platforms accountable for digital rights violations.**

The six objectives of this project are:
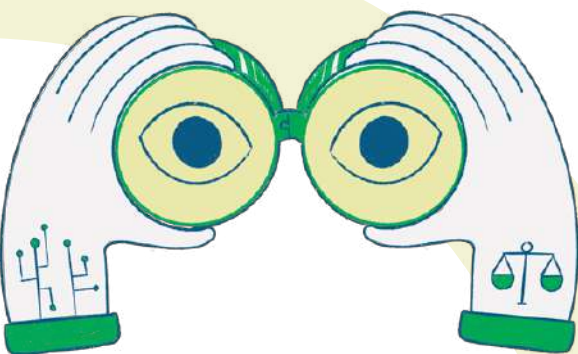
Align vision of tech research and advocates.

Identify where there is alignment in vision and goals between those leading tech-centred research projects and litigators/campaigners when it comes to platform accountability.

Explore litigation from tech research.

Explore opportunities for platform accountability litigation from the existing and ongoing work being carried out in tech-centred research projects.

Foster tech research and collaboration.

Foster and encourage greater collaboration and communication between tech-centred research projects and litigators/campaigners likely to pursue platform accountability cases.

Design protocols for evidence sharing/ storage.

Work toward collective design guidelines & protocols for the sharing, collection and storing of information, data and other evidence that can be leveraged for campaigning, advocacy, and litigation on platform accountability.
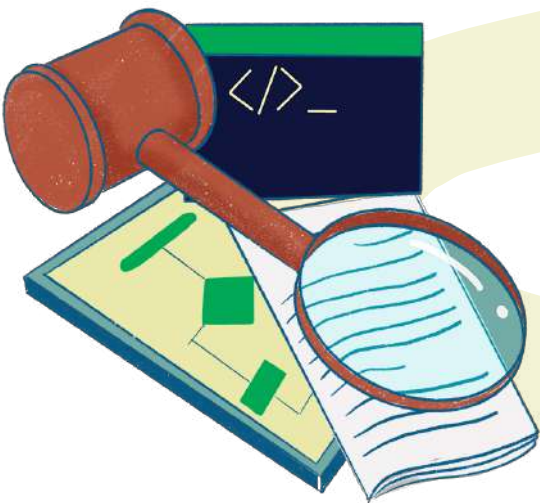
Find research & evidence for litigation.

Identify tech-centred research projects that are collecting information, data and other evidence that could be used in platform accountability litigation.

Expand research for advocacy and litigation.

Identify opportunities in existing tech-centred research projects to grow the scope of their work or improve working practices so their project outputs can be leveraged for campaigning, advocacy and litigation on platform accountability.

RESEARCH METHODOLOGY AND ORGANISATION OF RESULTS

To gather information about the current state of evidence gathering and potential solution areas, we convened a series of in person and online meetings. Much of our qualitative research presented in this report was gathered through two in-person meetings in October 2022, a capstone event in June 2023, and dozens of one-on-one calls, meetings, and discussions with members of the digital rights community including technical experts. The findings from all of these gatherings are presented in this report.

Readers should note that our methodology and research is grounded in the context of a European context. While DFF acknowledges the scope and reach of this work across continents, our specific organis ational focus is on assisting, strengthening, and supporting strategic litigation within the Council of Europe.

3 IN PERSON EVENTS

100+ PARTICIPANTS

20+ ORGANISATIONS INCLUDED

The April 2022 summit gathering consisted of mainly technologists in order to focus on what existing tech-centred research projects exist which could help drive platform accountability efforts. DFF hosted a first event in October 2022 to bring together litigators and activists to map what their needs for evidence were, and the gaps & challenges they faced in procuring that evidence. Some of the participants from the April summit were present in these meetings, in addition to litigators from DFF's network. The second meeting DFF hosted in late October 2022 brought those two groups together, in order to bridge the gap between them, and map out potential synergies and solutions. The initial qualitative insights from these meetings were compiled into a landscape analysis report that assessed the current state of the field, the needs & gaps of activists, and potential areas to build solutions.

During these meetings, we also discussed our common understanding of key terms such as platforms, platform accountability and digital rights. This was so as to ensure that different stakeholders including litigators, campaigners and technologists had a common framing. Our framing of key concepts in this report reflects the outcome of these discussions.

DEVELOPMENT OF EMERGING SOLUTIONS:

After the publication of the landscape analysis report in early 2023, personal feedback and one-on-one calls with stakeholders played a pivotal role in the developmental phase. Using the findings from the landscape analysis report as a springboard for discussion, we talked to members of the digital rights community such as researchers from various universities in Europe, like University College Dublin and the Technical University of Munich, civil society groups like Data & Society, as well as individual lawyers both in and out of the EU.

The individuals and organisations we spoke to in this phase of the project had not taken part in the initial workshops, and we used their participation and expertise to confirm that the challenges, insights, and solutions we had developed resonated with them as well. We also began discussing ways that those challenges could be addressed. Through discussion, outreach, and collaborative conversations, we started to align solutions with the identified gaps & challenges from earlier research.

FINAL CAPSTONE MEETING:

Building on the work from the past meetings and research, we hosted a final capstone meeting focused on sketching concrete solutions. Some tentative solution areas were then outlined and discussed by participants. The outcome of this meeting was an outlined set of solution areas and ideas for DFF and other organisations to pursue to concretely advance evidence gathering work.

DFF

THEMES

Capitalising on our shared vision for activism

In the field of platform accountability, litigators and technologists share a common vision of a more transparent, accountable, and just digital landscape, with both groups working towards greater insight into the inner workings of technology companies. **The solutions we propose capitalise on a shared vision to make the work of digital rights activism more effective.**

Opportunities for litigation & other channels of accountability

Technological exploration and legal advocacy are both driven by the desire to hold platforms accountable, but both are needed to enact change. Technological data, evidence, and strategies must be incorporated into the traditional legal process in a more substantial way than just plugging databases into existing cases.

Building & supporting a coalition of technologists & litigators

Lawyers have a wealth of legal expertise and knowledge about regulations and appropriate evidence procedures that are needed to develop and prove a legal case. While technology researchers have experience with best practices for data collection, storage, and analysis. **Therefore, a major theme of our insights and recommendations deals with the necessity of building this collaboration and community.**

Tools, Resources, & Evidence

Digital rights activists can utilise data breach reports, privacy policies, user agreements, and some existing information on platform algorithms as evidence to demonstrate the ways platforms violate digital rights. Existing website and investigative journalism techniques can be used as concrete ways for all stakeholders to expand their capacity to gather evidence.

DFF

INSIGHTS

In our discussion, participants worked together to identify needs and gaps when it comes to litigation and campaigning on platform accountability. For example, we thought through what information, data or evidence would strengthen current initiatives, and how different groups envision the future of platform accountability work.

With respect to our first theme, capitalising on our share vision for activism, we found an overwhelming uniformity in goals between all of the participants of our project. From lawyers working on case law to academics researching technology, everyone actively working on digital rights agreed that platforms 1) enable a suite of harms against digital rights, 2) these platforms lack the necessary transparency to be held fully accountable and 3) that gathering evidence about the inner workings of platforms is necessary.

Capitalising on our shared vision for activism

The digital rights world involves many different stakeholders, who pursue activism with different visions. One of our goals for this project was to speak to these stakeholders and identify where there is alignment in vision and goals between tech-centred research projects and non-technical activists like litigators and campaigners. We wanted to find ways that technical and non-technical activists could work together towards the shared goal of platform accountability, and to do so, we investigated the goals of each stakeholder group. Generally, litigators begin with case ideas, legal strategies and look for evidence that supports their legal hypotheses. On the opposite end, technologists more often begin by collecting data and forming their activist efforts around what evidence is available. However, the insights from this theme show that different stakeholders require different strategies, tools, protocols, and evidence for their different types of activism.

Overall, we found that **technologists and litigators align on the goal of holding platforms accountable.** However, there can be misalignment in strategies of activism and types of evidence gathering work needed. In our gatherings, we found that any toolkit for platform accountability would need to be flexible to the different strategies of activism pursued by different members of the digital rights field.

Similarities between technical and legal viewpoints exist. Over the course of multiple meetings that brought together litigators, campaigners, and technologists, we found that all used multiple types of evidence, focused on the context of each type of activism, and wanted more visibility into the existing tools and projects others have launched. Specifically, many practitioners wanted better communication and knowledge of existing projects instead of completely new tools or processes. There is a great deal of potential to capitalise on these similarities and work together.

When lawyers and technologists can leverage their shared vision, they can create effective projects. One of our participants and speakers at our June capstone event worked with PersonalData.io and HestiaLabs on a collaborative project between data scientists and lawyers where they used data donations and community organising of drivers to bring a case against Uber in Geneva.
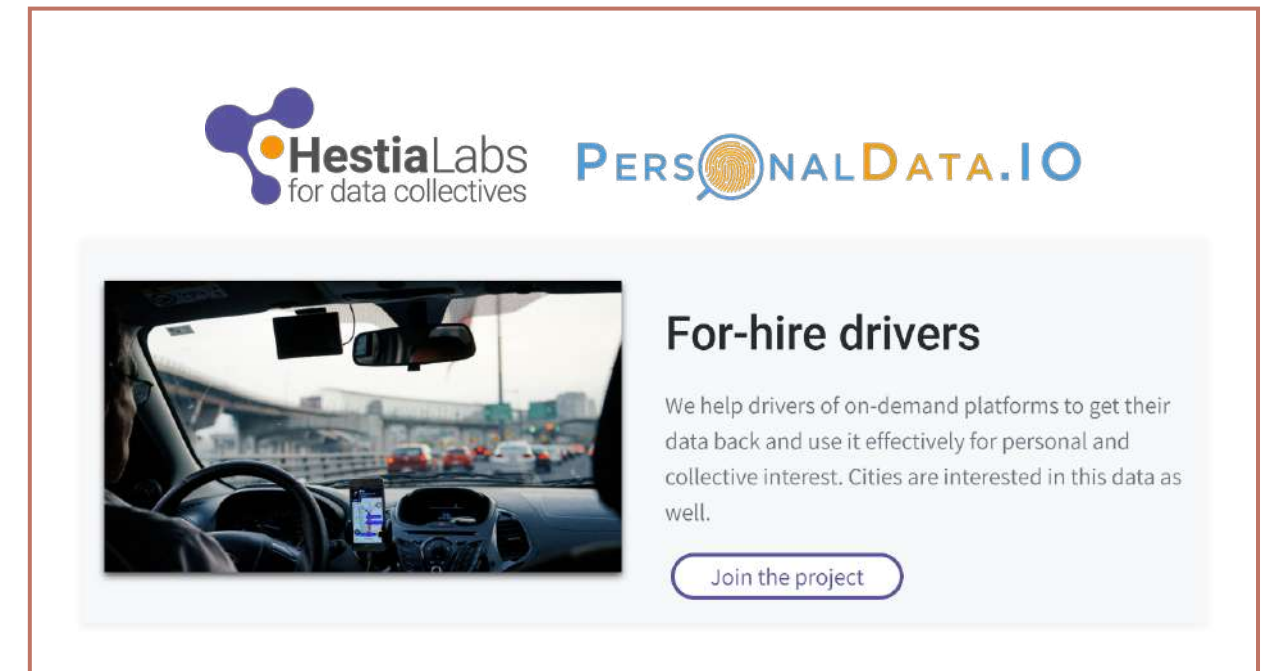
The trial focused on the right of access to data, and an initial decision was made earlier this year.

You can read more about this project at HestiaLabs' website, and more about the specific case they brough against Uber in this linked interview. In addition to their initial case in Amsterdam, HestiaLabs and PersonalData are now working with Uber drivers' collectives to generate a push for data donations. Working directly with these groups, they gather their data donations and format them into a readable, understandable user interface and excel document that can then be used by the drivers themselves and by lawyers taking Uber to court. By turning the original data donations into clear, readable evidence, they have made it easier for lawyers to understand the ways that Uber calculates ride distance, prices, and other internal metrics. As of the writing of this report, they have another ongoing case in Geneva to push for greater transparency into the Uber algorithm. This case shows the ways that different expertise (in law, technology, and data management) can be combined in service of the same shared goal – expanding digital rights for individuals (in this case for Uber drivers).

While there is a common thread of a shared vision among stakeholders, different strategies of activism require different evidence. **Holding platforms accountable can take many forms – advocating for affected communities, educating the public through campaigns, and direct litigation are just a**

few examples. Fighting for digital rights requires many different stakeholders and activist groups – campaigners, strategic litigators, legal researchers, and more. In our conversations around platform accountability, participants used different types of evidence for different channels of activism. Campaigners wanted to hear stories from affected communities, so that they could use those stories in public education efforts, ongoing campaigns for policy, and campaigns for additional regulation of technology companies. Litigators submitting complaints and cases to regulatory bodies in the EU found that having knowledge of internal documents or procedures of technology companies proved helpful. One participant was able to use internal documents they had uncovered to tell regulators what companies to look at for violations of data privacy, where to look, and what specifically those regulators needed to ask for. Even within the specific field of litigation, the types of evidence activists need differ between case law and pre-trial research. In the HestiaLabs and PersonalData case study, the team understood that the specific data they needed was data donations from Uber drivers. However, other teams might need anecdotal stories from affected communities, whistleblower information about inside company practices, or some other type of evidence about the platform.

Opportunities for litigation & other channels of accountability

One of the major channels by which we can hold platforms accountable is direct strategic litigation. In the European context, especially within the EU, there's a promising space for strategic litigation against tech giants. For instance, the GDPR can be used to force companies to provide users with their data, so there is a potential for data collection through data donation drives. Other new potential avenues for work include the EU Charter and Collective Redress Directive. While much litigation is not yet finalising, there is a great deal of potential in exploring litigation through the DSA, DMA, and upcoming frameworks such as the AI Act. Within these legal contexts, evidence becomes a powerful tool for shaping legal discussions, setting precedents, and strengthening legal cases.

However, litigators currently face major challenges to gathering evidence for platform accountability cases, namely a **lack of resources** faced by both litigators and their technical counterparts, **gaps in knowledge** about the basic technical components of the algorithms and data practices used by platforms, and an overarching lack of strategy around evidence gathering since most cases gather evidence on a case-by-case basis.

Lack of resources

Many litigators encounter serious challenges in gathering technical evidence within the context of platform accountability work. Not only does the technical complexity involved in some of the data downloads, scrapes, or information searches exceed their knowledge, but they also do not have the deep technical knowledge to know where to look for that data in the first place. The baseline difficulty of obtaining any data from opaque platforms compounds the difficulty of finding and analysing data.

As many of our participants told us, litigators are deterred from pursuing projects that require technical expertise due to the formidable hurdles they face in navigating the intricate technological landscape. Not necessarily because they lack the ability to develop the technical knowledge but because they lack the time and resources to acquire that knowledge.

Gaps in litigators' technical skills

While our project had many participants, many of them were litigators working on traditional litigation. However, understanding and using basic technical ideas is something that litigators are capable of, and they need to overcome this when working on digital issues.

For example, a common technical process in the data world is ETL (extract, transform, load). When bringing data from multiple sources, data scientists have to extract that data from its original source (an image database, for example), transformed (like removing all duplicate images), and loaded into a new database. Imagine gathering evidence from multiple witnesses, documents, and sources (Extract), then meticulously reviewing, organising, and possibly translating this information into a unified format (Transform), and ultimately presenting it in a well-structured manner to the court (Load). This method ensures that the data is accurate, coherent, and optimized for analysis, much like the way legal evidence is prepared for courtroom presentation.18 Another tool used by many technical researchers are Application Programming Interfaces (APIs) which can "scrape" information from different internet sites and databases, which involve extracting information from online sources by interacting with APIs.19 Envision requesting specific documents from various government agencies using proper protocols. There are many legal channels that lawyers use to extract information through existing channels, like GDPR or Freedom of Information requests. Similarly, API scraping entails programmatically accessing online platforms' data through designated interfaces, rather than manual browsing, to collect data for legal analysis. Just as a lawyer follows regulations when obtaining official records, API scrapes adhere to website terms of use and relevant laws while extracting data for purposes such as legal research or market analysis. Unfortunately, many litigators see terms like ETL and API as outside of their scope of abilities, and fail to learn the necessary skills, protocols, and tools to recover, manage, and learn from data.

Gap in knowledge

Even when litigators do have the chance to work with data donations, data pulled from data requests, or other forms of technical evidence, many lack the basic technical knowledge to be able to put that evidence in context. Without understanding the basics of data usage within algorithmic systems, litigators might be unable to use the technical evidence they have. On the other hand, many technologists lack basic knowledge of the law and the specific litigation process and could benefit from a basic primer on how their data-centric projects can be leveraged to pursue platform accountability cases. For instance, litigators could share with technologists more insights into how evidence should be presented to non-technical court benches so that the judges making the decisions on these cases can make sense of the evidence presented and use it to decide against platforms in platform accountability cases. Also, a basic knowledge transfer from litigators to technologists on basic legal proceedings including the admissibility of evidence and burdens of proof could go a long way to ensure that evidence developed by technologists can be effectively used in court proceedings.

Many organisations do not have the time, human resources, technical expertise, time, or money to do technical evidence gathering, regardless of the availability of the data, and therefore do not end up pursuing technical projects.

Case-by-case basis for evidence gathering

Litigators we spoke to generally agreed that they gathered evidence for their work on a case-by-case basis. As cases emerged, they would seek out specific evidence that fit into the context of that case. However, few litigators had developed a robust data collection strategy that could be used across cases.

Litigators shared that their data practices too often fell into the same jurisdictional silos that their other legal work falls into. Within the EU especially, a case may be brought on similar grounds and dealing with similar issues in one jurisdiction by one organisation to another case being brought by another organisation in a different jurisdiction. These cases may require the same data collection and research but both organisations are independently collecting and collating the data. This is a waste of the limited resources available to civil society.

According to technical researchers and investigative data journalists, having a transferrable, repeatable data collection strategy & protocol is an important part of understanding technical evidence. While litigators' case-by-case basis is currently the norm, it is not be the most effective way for them to collect and use technical evidence.

Building & supporting a coalition of technologists & litigators

A recurring theme throughout our workshops and events is the need for litigators to share resources, knowledge, and educational materials within a trusted context. Litigators want to be able to speak to each other and to technologists in a forum where they can trust that other participants are well-intentioned, knowledgeable, and discreet. Furthermore, they want assurance that their communications and work will be secure, safe, and unmonitored. Such a community of practice would enable them to more effectively litigate and campaign on issues of digital rights and platform accountability.

Building a trusted community of practice consisting of litigators and technologists would help litigators incorporate better strategies of technical evidence gathering, avoid duplicating the work of others, and assist in building the capacities of both litigators and technologists through sharing their respective expertise. Strengthening the community of individuals and organisations that work on digital rights feeds into the creation of outputs like legislation, regulation, policies, and practices. While building a community of practice would be valuable alone, it would also feed into collaborative litigation projects led by community members. Ultimately, the outcome of our community would include litigation that advances and promotes digital rights, leading to positive changes that yield benefits across society, but especially for the groups, communities and individuals most impacted by technological harm and digital rights violations.

Part of capitalising on technologists and litigators shared visions for activism **involves creating a forum for discussion, sharing expertise, and connecting on potential projects.** In all the participatory workshops and qualitative research we have conducted in the past year, the lack of such a forum represents a clear gap, and a need that could be filled with future work. The digital rights community operates on a system of interpersonal trust and communication. Although many litigators want to work more closely with technologists to litigate against platforms more effectively, they need to trust those technologists first. One way to establish trust is to create a vetted, curated community of activists who can share their expertise while also trusting that their collaborators share the same goals as they do. Building such a community of practice would enable them to more effectively litigate and campaign on issues of digital rights and platform accountability.

Furthermore, many of the other emerging areas of evidence gathering work would greatly benefit from the existence of a shared community. For example, litigators and technologists both expressed the necessity of **shared community protocols, data collection strategies, and tool lists,** which could all be more easily sourced from a trusted community group.

DFF

Tools, Resources, & Evidence

Gathering evidence requires technical artifacts, tools, and data collection strategies.

Technical data gathering is important, and many digital rights activists acknowledged the importance of analysing data donations, establishing robust data-sharing protocols, and managing digital data effectively in the process of litigation.

However, litigators and technologists use different mechanisms to gather evidence, and not all evidence on platforms is necessarily technical in nature. Pieces of information that would be helpful for litigators are often obtained through processes, like freedom of information requests. Technical expertise would be useful in establishing standards of data handling, and protocol for data donations.

Many tools already exist, and simply need to be maintained or publicised. For example, OCCRP ALEPH is an online tool already used by investigative journalists to dissect public datasets and information about public entities and countries. Other tools include the Department of Secrets wiki page for whistleblower documents and other leaked information, and an ongoing project by HURIDOCS to build out a flexible database application designed for human rights defenders to capture and organise collections of information.
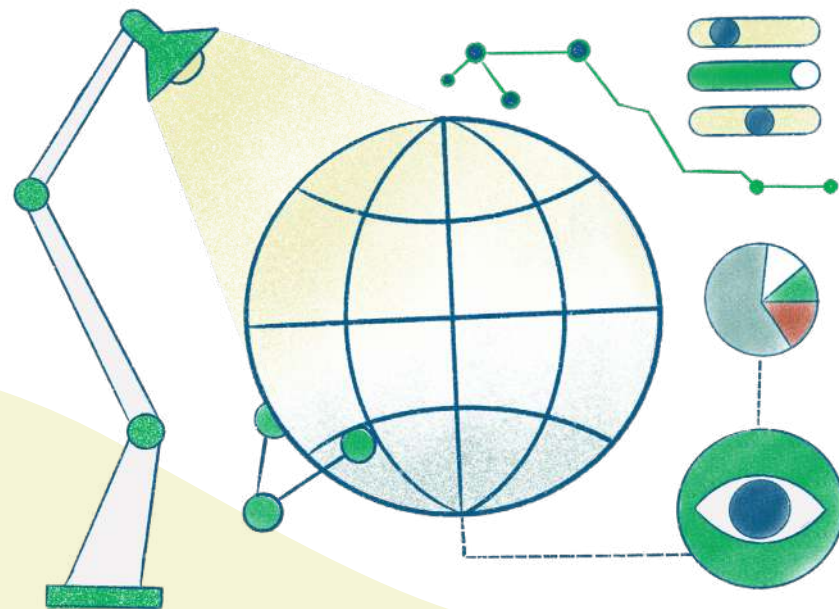
Many of the organisations we worked with during this project have created tools and projects of their own. For example, Tracking Exposed has created many tools and projects, specifically for platform accountability, including DFF-funded litigation. Another example is AlgorithmWatch, which has many projects related to auditing algorithms for systemic risks, or the DataSkop platform for data donations, among others. Mnemonic's Digital Evidence Workflow based on the Electronic Discovery Reference Model developed by Duke University School of Law. These tools, and several dozen others uncovered in our work this past year, are included in Appendix A of this report. While this list is certainly not exhaustive, we hope that it shows how many ongoing projects, websites, and reports currently exist to address the issue of gathering evidence against platforms. Although more can always be done, centralizing and publicising these tools would mark an important first step.

Because many of the litigators, technologists, and non-profit workers in the digital rights space have no central repository for existing tools and information, they often end up duplicating each other's efforts, wasting everybody's limited time. **Duplication of effort is a real problem that our participants faced,** and many suggested that better dissemination of existing tools would be the best way to address this problem. Instead of just creating new tools, activists need better insight into the ones that already exist. **Participants expressed a greater need for funding, maintenance, and dissemination of current projects.**

RECOMMENDATIONS

Capitalising on our shared vision for activism

Collaborative knowledge hub **01**

02 Non-legal remediation efforts

Developing better data collection strategies **03**

04 Technical data pipeline

Technical education for litigators **05**

06 Community of practice

Tools, Resources, & Evidence **07**

1. Collaborative knowledge hub

The digital rights community contains an astonishing amount of collective knowledge and experience about platform accountability, effective litigation, and embodied knowledge on the inner workings of many platforms' algorithms. During the three in-person events DFF hosted as part of this project, much of the discussion centred around ongoing projects, practices, and inside knowledge about the best ways to engage large platforms in and out of court. Some of our participants had stories to share about the judicial culture of certain courts, or what types of information specific EU regulatory agencies were most interested in. Some knew about different ways to request information, the best types of language to use in freedom of information requests, and how to get gig workers to show up to an education session about data privacy.

Recognizing the wealth of expertise that resides within this community, one of the most promising areas of future work would involve the creation of a comprehensive knowledge hub for platform accountability work. This might look like a digital repository, a shared online folder, or a forum for posting and organizing information. Such a hub would not only facilitate the dissemination of valuable information but also foster collaboration between members of the digital rights community.

Rationale for the Collective Knowledge Hub:

As we discussed above, there is already a shared vision and passion among many of the litigators, technologists, and civil society professionals working on digital rights. Shared information between different experts has been a successful way to advance digital rights – like in the example of PersonalData and HestiaLabs. Fostering more of this type of interaction and knowledge sharing is not only something that many people want, but an initiative that would pay real dividends in making our collective work easier and more effective.

What the solution entails:

Building out a basic version of a knowledge hub could be as simple as setting up an online forum page with moderators' self-organizing information, or as complex as creating a secure, data-enabled server to share sensitive documents, data, and information.

In increasing order of complexity, central knowledge hub artifacts could include:

An email thread, hosted by DFF or another organization. One group could serve as the aggregator of knowledge, and community members could respond to a newsletter or other email communication with information, updates, and shared knowledge to be sent out in the next newsletter.

An online forum page, like the Coalition for Independent Technology Research (listed in Appendix A as one of the tools we identified). On this site, litigators could upload question in the form of public posts, ask questions of the broader community, and ask questions of others.

An online interactive database, where not only information, but also datasets and documents could be shared securely. A site like this would be of the same order of complexity as something like Distributed Denial of Secrets, a whistleblower information site. It could be hosted publicly or privately.

Stakeholders in solution creation:

Creating a central repository for information and knowledge on evidence gathering would be a collective effort. DFF could help support initial versions of the project and use our grantee community to build it, but ultimately a collaborative knowledge hub would need to be built and maintained by more than just our organisation. Any future solution would need to consider how to fund and maintain this solution sustainably. Potential avenues might include collaborative efforts by more than one organization, rotating funding for maintenance, or partnerships with universities.

While this future work can start small and grow to become more complex, it is important that we start collecting and collating our shared knowledge. Bringing together our collective wisdom is an important step in capitalising on the shared vision of technologists and litigators.

2. Non-legal remediation efforts

Developing platform accountability work also means moving past pure litigation and pure technical research as methods of gathering evidence. It's important to acknowledge the varied constellation of activist efforts that could also be supported moving forward around platform accountability including community organising and labour organising groups, and activism that centres and supports affected and marginalised communities.

Rationale for the non-law remediation efforts:

There are many different types of people and organisations who fall under the broad term "digital rights activist." While our participants were mainly involved in pure litigation work, we also worked with campaigners, educators, community organisers, and workers' rights groups. Furthermore, there are many different activists beyond even the groups that DFF worked with for this project who focus on non-law activism efforts. Throughout the course of this project, we found that lawyers could greatly benefit from technologists' perspective. Furthermore, litigators and technologists can learn from activists who focus on non-legal efforts like education, campaigning and community organising.

What the solution entails:

Examples of non-law digital rights work already exist. For example, the project Terms-we-Serve-with is a "feminist-inspired multi-stakeholder engagement framework and tools for evolving and enacting social, computational, and legal agreements that govern the lifecycle of an AI system." This project centres not around litigation, but around the actions of data subjects and users – ensuring they can refuse data-sharing, which involves them being informed about data-sharing processes. It also centres around the concept of contestability, the ability for people to disagree, challenge, dispute, or otherwise express their concerns. Crucially, these mechanisms do not revolve around direct litigation, but user empowerment.

Finally, community mobilisation efforts can assist in evidence gathering for platform accountability. Helping to educate groups affected by platform-based harms will help litigators better understand their needs, the context they operate in, and their lived knowledge of how platforms propagate harm.

DFF

Stakeholders in solution creation:

Various organisations and individuals are actively engaged in non-law remediation efforts, to address social and systemic issues. Educational initiatives aim to raise awareness, empower marginalised communities, and foster critical thinking around pertinent issues, like the work being done in the UK by the Public Law project that focuses on educating the public about how their data is used. Community organising brings together residents, activists, and stakeholders to collectively advocate for change and amplify marginalised voices. These multifaceted approaches underscore a growing recognition that sustainable and inclusive progress requires collaborative efforts that extend beyond legal frameworks.

Opportunities for litigation & other channels of accountability

3. Developing better data collection strategies

As we discovered, many litigators currently seek evidence on a case-by-case basis, and lack a comprehensive, repeatable data collection strategy. To effectively gather evidence about the platforms our community is working to hold accountable, we need to address the issue of case-by-case collection of data and develop a technical data collection strategy. Further work in this area could involve developing a data pipeline, shared database, or online platform to share data.

During the DFF Evidence Gathering Capstone held in Berlin in June 2023, an investigative data journalist presented on the crucial topic of data management for investigative journalism and its potential application in strategic litigation. Crucially, data journalism requires a robust data management strategy that is included from the very beginning of a project i.e. from the very inception of a story idea. Litigation, while different from journalism, requires a similar level of strategic thinking about data acquisition.

Rationale for data collection strategies:

Case-by-case evidence gathering is not only ineffective, but usually goes against best practices for data collection and data science. Not only would better data collection strategies help litigators, but these strategies largely already exist, or can be adapted from data scientists and investigative data journalists. Creating a broad data strategy for litigators would be an easy first step in educating the community about how to best gather, collect, and use data in their work.

What the solution entails:

Strategies should include structured and standardised data management protocols that encompass methods for data acquisition, storage, organization, and analysis, while prioritising data security and privacy. Many litigators we spoke to agreed that data collection was a challenge, because they often only started looking for data when they needed it for an existing case or argument. By adopting a more comprehensive and proactive data management practice, litigators can strengthen their ability to gather, organize, and present evidence for their cases on platform accountability.

An initial project in developing data strategies for litigation could be broad and borrowed from other groups. However, data must be understood within its context, so a comprehensive data strategy for litigators would need to be developed by a working group involving lawyers, campaigners, and other stakeholders within the digital rights community. Such a framework could be used as a starting point by litigators who want to incorporate evidence gathering and data collection into their cases.

Stakeholders in solution creation:

Technical participants in DFF's evidence gathering project have included data journalists, security engineers, data scientists, and data management professionals. These technical professionals know that data strategy is already a topic in the technical world, and that best practices for gathering, storing, and disseminating data responsibly already exist. Their expertise could be combined with that of litigators to form a data strategy for litigation.

4. Technical data pipeline

Even with a good data collection strategy, data needs to be correctly handled and processed to be of any use. As we have discussed in our insights and themes, many litigators have good data, but even if they manage to use it themselves, they end up not using it after their cases. Our participants voiced their frustrations with collecting data only to find that someone else had already gathered it, replicating work done by other researchers.

In order to solve this technical data-sharing problem, a future solution might involve the creation of a technical data pipeline. Such a pipeline would involve a structured and automated process for collecting, processing, and analysing relevant legal data to support litigation efforts. This pipeline would start with data extraction from various sources, such as legal documents, court records, and regulatory databases. The data could also be uploaded by technical researchers, academics, or other litigators. Part of a technical data-sharing toolkit could involve transformations to ensure consistency, accuracy, and compliance with principles and laws like the GDPR. Once prepared, the data would be loaded into a secure and accessible database or data warehouse, making it easier for litigators to search, retrieve, and analyse information efficiently.

Rationale for a technical data pipeline:

Data from many different sources exists that could be used in platform accountability litigation – data donations from Uber drivers, whistleblower commentary on internal Facebook algorithms, information scraped from Twitter's API before it was shut down by Elon Musk, and surely many more places.20 This data is in different forms, and requires different levels of security, storage, and data handling. However, it is possible to create a centralized data pipeline that could be shared among activists in the digital rights space. While it would require investment and technical development, technologists could help create a shared data hub where litigators, campaigners, and researchers could go to seek evidence supporting their own projects. Building such a data pipeline would make data collection much easier for many litigators.

What the solution entails:

A digital rights data pipeline might simply be a secure excel-sharing mechanism on a website. However, a more cutting-edge solution would involve a forum for the extraction, aggregation, and sharing of data, a pipeline that could be responsive to privacy regulation and contextual standards for different countries within the Council of Europe. Data from different projects, cases, and campaigns could be securely loaded into a centralized, open source, cloud-based repository, accessible to authorised users. A pipeline like this could not only streamlines research efforts but also facilitates collaborative case preparation and give litigators more data to shape their arguments in court.

Creating a technical data-sharing mechanism or a data collection strategy would certainly face challenges, such as addressing privacy concerns, ensuring usability and user design, and centring affected populations in the data collection process. However, addressing the technical issues of data collection and storage through the development of a data collection strategy and/or the creation of a data sharing pipeline for the digital rights community would be huge contributions to the field of platform accountability.

### Stakeholders in solution creation:

Participants at our events included members of technical organisations that focused on building data pipelines and ensuring technical security of data. In order for DFF or others in the digital rights space to successfully build out a technical data pipeline, we would need to work with technical groups like these.

5. Technical education for litigators

Legal practitioners face an increasing need to grapple with complex technological issues in digital rights cases. Unfortunately, many lawyers do not possess the technical know-how to understand data processing, machine learning, and other critical technological aspects relevant to evidence gathering. This knowledge gap often results in missed opportunities, incomplete arguments, and potentially unsuccessful litigation.

DFF

Rationale for technical education:

A simple solution to many of the issues litigators face in evidence gathering is to just educate them about the ways technologists already gather evidence. As we found in our project this past year, many litigators shy away from terms like ETL or API because they seem technical and alien. In reality, litigators could likely learn to use these tools to their advantage within a fairly short time frame. Therefore, it would be beneficial to build out a program that focuses on educating litigators and non-technical digital rights actors in order to empower them to collect data-based evidence for their work.
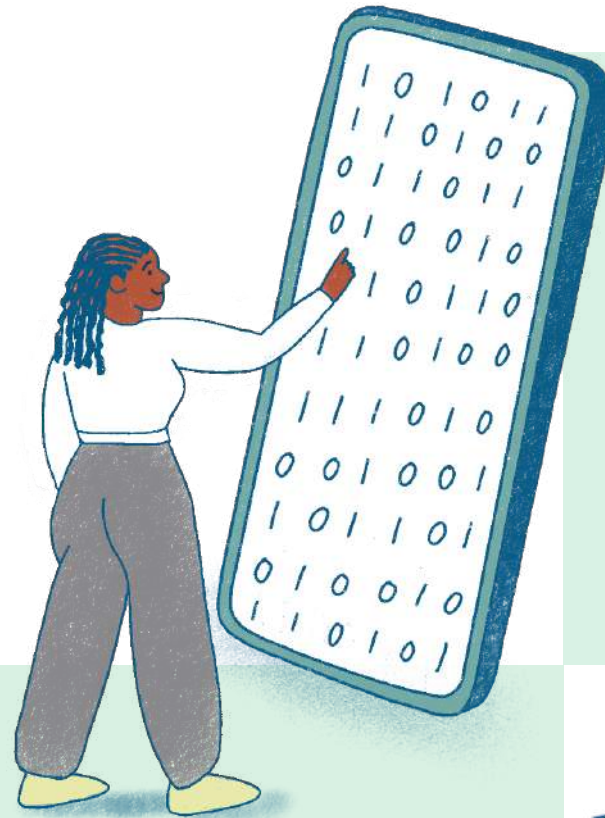
Technical expertise: As mentioned in the discussion of the technical data pipeline, DFF has collaborated with some purely technical organisations in the course of this project. With their help, we could host expert sessions on more advanced topics like data pipelining. Other expertise could include security experts, hacking for good, or data strategy professionals.

There are many different opportunities for educational solutions to evidence gathering, and they could be set up in a multitude of ways. For example, participants could provide input on how education is structured, different organisations could provide different jurisdictional information, or individual experts could contribute their knowledge in the form of simple lesson plans. Regardless, addressing the knowledge gap in this area is an important step.

What the solution entails:

Education for litigators about technical evidence gathering could take many forms. Like previous suggested solutions, these forms could have various forms of complexity. From least to most complex, here are some artifacts that DFF or others could invest in creating:

Educational Newsletter: A simply email chain with basic information on APIs, ETL, machine learning techniques, and common data processing could help litigators understand more about gathering technical evidence.

Technical Course: After asking litigators what information they feel they need and assessing the gaps in their knowledge, we could create a very basic course covering technical evidence gathering. This course could be a simple series of white papers, or more interactive videos. Some of the learning resources that this would include could be a basic understanding of how machine learning works, or a basic algorithmic understanding course. One lesson might cover Natural Language Processing 101, or how to use an API to scrape data. These lessons could be created by members of our own community, including many of the people who participated in this project. Many courses like this already exist, so our community could leverage existing resources and classes online or in person in support of this goal.

Integrating Case Studies: A more thorough educational deliverable could involve case studies and information from successful litigation. Groups like Panoptykon who have used technical evidence in their cases could talk about what they learned and share key takeaways in an educational format.

Stakeholders in solution creation:

Several groups we spoke have already developed educational materials that litigators could benefit from, like Ethical Intelligence and HestiaLabs. If DFF pursued an education-focused solution, we could use these resources to develop a program tailored to our community of litigators. DFF could also identify technologists within the digital rights community to help build out a contextual educational program.

Building & supporting a coalition of technologists & litigators

6. Community of practice

So much of the work we did this past year involved participation from the digital rights community, and much of the feedback from our participants revolved around the need for stronger community ties and more help from others. Building a trusted community of practice consisting of litigators and technologists would help litigators incorporate better strategies of technical evidence gathering, avoid duplicating the work of others, and assist in building the capacities of both litigators and technologists through sharing their respective expertise.

Rationale for a community of practice:

Strengthening the community of individuals and organisations that work on digital rights feeds into the creation of outputs like legislation, regulation, policies, and practices would support all of our other proposed solutions, from building educational materials to creating a technical data pipeline. A community of practice between litigators and technologist would be valuable in and of itself, but within the context of evidence gathering it also serves to support all of our other recommendations and proposed solutions.

Stakeholders in solution creation:

Broadly, a community of practice would include all stakeholders working on evidence gathering for platform accountability. More realistically, this community would likely include DFF's network, grantees, and colleagues within the digital rights space, and would have to expand to include technical researchers, academics, and data professionals.

What the solution entails:
Our proposal for building a trusted community of practice for technologists and litigators has three main components:

1. Establishing recurring space for discussion: Host a dedicated platform for ongoing discussions among litigators, providing a secure and trusted space for knowledge sharing, resource dissemination, and dialogue on technical aspects of litigation. Creating a space for these discussions is an important component to community building, as it provides a forum to share accomplishments, advice, and evidence.

2. Providing matching for individuals and organisations: Facilitate connections between our grantees and other community members, and experienced technologists. Encourage strategic sessions and foster collaborations to enhance the technical capacity of litigators by establishing their technical evidence gathering strategy early in their process and connecting them with technical professionals.

3. Creating a forum to promote work: Establish a forum where digital rights practitioners can showcase their work, share best practices, and contribute to a living document listing ongoing tools, projects, and resources in technical evidence gathering. Our research on evidence gathering has uncovered existing projects and toolkits that we could disseminate more widely to the community. While a forum like this would be informal and less searchable than something like the collaborative knowledge hub we proposed above, it would be a helpful way to quickly share information and ask colleagues for help.

While many of our other solutions recommend more specific outcomes, building a community of practice is an overarching goal. It would support other efforts to build our capacity for evidence gathering, but it would also be a natural side-product of our other proposed work. Implementing our other solutions depends on a group of people working together to provide technical expertise and legal knowledge, so a community of practice complements the other recommendations in this report.

7. Tools, Resources, & Evidence

During this project, many of our participants have shared tools, websites, resources, and sources for evidence from their own work and experience. An initial list of these tools has been attached to this report, but more comprehensive work could be done to catalogue the existing suite of tools and projects that support evidence gathering work. As many of our participants noted, there is a great deal of duplication of effort in the field today that could be eliminated through better sharing of tools. Furthermore, with a clearer picture of the existing tools in the field, funders could focus on providing funding for the maintenance of helpful tools, instead of devoting resources to creating new tools that duplicate existing efforts.

For example, the online tool "Have I Been Trained?" is a search function developed by a coalition of artists and technologists that will tell you if your photo or art has been used in the training data sets for any generative AI model that makes photos or art. DALLE-2 from OpenAI might have used your picture in its training data for images, and this tool can tell you that. If someone wanted to pursue litigation against one of these companies for breaking copyright law regarding the use of their art, this tool would serve as an excellent tool for providing evidence.

While our tool list included in this report represents a start, more work can be done to formalize the sharing of tools, resources, and technical evidence sources among members of the digital rights community.

CONCLUSION

A year ago, in August of 2022, we started with a set of six objectives that revolved around developing the digital rights community's capacity for gathering evidence to hold large technology companies accountable. After three in-person events and many online phone calls, roundtables, and feedback sessions, we found that technologists and litigators broadly share the same vision but need to develop stronger bonds to be effective in their work. These bonds can be social, through sharing information, knowledge, and resources, or technical through the creation of a data pipeline, technical educational materials, or sharing data itself.

The findings of our year-long investigation shed light on the crucial role evidence gathering plays in fostering platform accountability amidst the opacity of large technology companies. Our recommendations imagine a future digital rights landscape where knowledge is shared openly and effectively, and a collaborative knowledge hub and a community of practice can facilitate the exchange of insights and strategies, and avoid wasteful duplication. Additionally, we advocate for concrete technical solutions, such as a data sharing pipeline that empowers litigators to share vital information and evidence. We have also uncovered a range of existing tools and resources that we started mapping and categorising.

In conclusion, our exploration of evidence gathering's role in platform accountability underscores the need for collective action, strategic innovation, and interdisciplinary collaboration. By implementing these recommendations, we can pave the way towards a more transparent, just, and accountable platform economy and digital ecosystem for all.

DFF

CITATIONS

1. Taylor, L. (2017). **What is data justice?** The case for connecting digital rights and freedoms globally. Big Data & Society, 4(2), 2053951717736335.

2. Klang, Mathias; Murray, Andrew. 2005. **Human Rights in the Digital Age.**

3. United Nations. (2017). **Ensuring the protection of human rights in the digital era. Office of the Secretary-General's Envoy on Technology.** http://www.internetrights.org.uk/index.shtml?AA_SL_Session=-8fa795873994ed10dd54938b98227a99&x=609

4. Zuboff, S., (2015). **Big other: surveillance capitalism and the prospects of an information civilization.** Journal of information technology, 30(1), pp.75-89.

5. Johansson, A. (2019, September 6). **How Startups Develop and Deploy Matching Algorithms. Entrepreneur.** https://www.entrepreneur.com/science-technology/how-startups-develop-and-deploy-matching-algorithms/338442

6. Kastrenakes, J., & Vincent, J. (2023, June 26). **Hope, fear, and AI. The Verge.** https://www.theverge.com/c/23753704/ai-chatgpt-data-survey-research#:~:text=One%20finding%20is%20particularly%20clear,category%20within%20these%20creative%20fields.

7. Singh, S. (2019, July 22). **Holding Platforms Accountable: Online Speech in the Age of Algorithms. New America.** https://www.newamerica.org/oti/reports/report-series-content-shaping-modern-era/

8. Justice Department Secures Groundbreaking Settlement Agreement with Meta Platforms, Formerly Known as Facebook, to Resolve Allegations of Discriminatory Advertising. (2019). **Department of Justice | Homepage | United States Department of Justice.** https://www.justice.gov/opa/pr/justice-department-secures-groundbreaking-settlement-agreement-meta-platforms-formerly-known#:~:text=The%20proposed%20agreement%20resolves%20a,familial%20status%20and%20national%20origin.

9. Mac, R., & Kang, C. (2021, October 3). **Whistle-Blower Says Facebook 'Chooses Profits Over Safety'. The New York Times.** https://www.nytimes.com/2021/10/03/technology/whistle-blower-facebook-frances-haugen.html

10. Despite flexibility, gig work and insecure income are harmful to U.S. workers, according to UTHealth School of Public Health. (2022, August 22). **Houston Campus - Campuses - UTHealth Houston School of Public Health.** https://sph.uth.edu/news/story/despite-flexibility-gig-work-and-insecure-income-are-harmful-to-us-workers-according-to-uthealth-school-of-public-health & Weil, D. (2023, January 26). What's a 'Gig' Job? How it's Legally Defined Affects Workers' Rights and Protections. SHRM - The Voice of All Things Work. https://www.shrm.org/executive/resources/pages/defining-gig-jobs-2023.aspx#:~:text=As%20independent%20contractors,%20workers%20lose,now%20required%20in%20many%20states.

11. Weil, D. (2023, January 26). **What's a 'Gig' Job? How it's Legally Defined Affects Workers' Rights and Protections. SHRM - The Voice of All Things Work.** https://www.shrm.org/executive/resources/pages/defining-gig-jobs-2023.aspx#:~:text=As%20independent%20contractors,%20workers%20lose,now%20required%20in%20many%20states.

12. Goldstein, J., & Grossman, S. (2021, January 4). **How disinformation evolved in 2020. Brookings.** https://www.brookings.edu/articles/how-disinformation-evolved-in-2020/

13. Gonzalez, J. (n.d.). **Platform Accountability. Free Press.** https://www.freepress.net/issues/media-platform-accountability/platform-accountability

14. Vakkuri, V., Kemell, K. K., & Abrahamsson, P. (2019). **AI ethics in industry: a research framework. arXiv preprint arXiv:1910.12695. What's Wrong with Transparency Reporting (and How to Fix It). (2021). ADL.** https://www.adl.org/resources/report/whats-wrong-transparency-reporting-and-how-fix-it

15. Maggiolino, M. (2019). **EU trade secrets law and algorithmic transparency.** Available at SSRN 3363178.

DFF

16. Fujiwara, T., Muller, K., & Schwarz, C. (2023, July). **The Effect of Social Media on Elections: Evidence from the United States.** https://www.princeton.edu/~fujiwara/papers/SocialMediaAndElections.pdf

17. Allcott, H., & Gentzkow, M. (2017). **Social media and fake news in the 2016 election.** Journal of economic perspectives, 31(2), 211-236.

18. Costello, T., Blackshear, L., Costello, T., & Blackshear, L. (2020). **What Is ETL?.** Prepare Your Data for Tableau: A Practical Guide to the Tableau Data Prep Tool, 1-3.

19. Trezza, D. (2023). **To scrape or not to scrape, this is dilemma.** The post-API scenario and implications on digital research. Frontiers in Sociology, 8, 1145038.

20. Bell, K. (2023, April 6). **Twitter shut off its free API and it's breaking a lot of apps | Engadget. Engadget.** https://www.engadget.com/twitter-shut-off-its-free-api-and-its-breaking-a-lot-of-apps-222011637.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAM4bUfHgDoyM_RgQFjKJw8QYPI-dde7M2GDJeUQPl8RYwFL9hg5xSTQ_DvtptwfPmweUquu2OCIuDpel-mH7eEZ0tCy7eBIISwnKk_U_46eJ5stTtw6d3XlXgEB10_4Sxb7UH-xItSydeyhm-93TpBQH63gza26UxVUAUNkpie6480F

APPENDIX A:
Evidence gathering tool & resource list

ORGANIZATION	DESCRIPTION	LINK
Brave New Europe	Project compiling information from members of the gig worker community	https://braveneweurope.com/the-gig-economy-project
Coalition for Independent Technology Research	Group of researchers, technologists, and activsts with a popular online forum for sharing information and data	https://independent-techresearch.org/
Connected by Data	List of participatory data governance practices in use globally	https://connectedbydata.org/cases
Correctiv	Court donations records	https://corrctiv.org/spendengerichte
Crowd News Room	Investigative journalism resources	crowdnewsroom.org
Data Europa EU	Access to 1.4 million public datasets from 36 countries (EU, EEA, Switzerland and EU Neighbourhood states).	https://data.europa.eu/en
Digital Public Goods	A digital public goods registry of tools that have been reviewed	https://digitalpublic-goods.net/registry/
DigiWhist	Cutting-edge sources of open data readily available in a structured downloadable format.	https://digiwhist.eu/resources/data/
Distributed Denial of Secrets	Whistleblower site for news leaks	https://ddosecrets.news/index.php/Main_Page
EU Open Data Portal	Open data from EU institutions, agencies and other bodies	https://data.europarl.europa.eu/en/home
Follow The Grant	Tracing of grant funding	followthegrant.org
Github	List of publicly available APIs that can be used to scrape data and gather evidence online	https://github.com/public-apis/public-apis

ORGANIZATION	DESCRIPTION	LINK
Glaze	A tool that disrupts generative AI through image creation	https://glaze.cs.uchicago.edu/
Have I been trained?	Online tool made by an artist coalition to see if artwork or photos have been used to train generative AI	https://haveibeentrained.com/
HURIDOCS Uwazi	A flexible database application designed for human rights defenders to capture and organise collections of information	https://uwazi.io
ICIJ Datashare	Software used to organize investigative data	https://datashare.icij.org
Model	Investment and money tracing database	http://followthemoney.tech
OCCRP Aleph	A global archive of research material for investigative reporting containing leaks and other data	aleph.occrp.org
OpenSecurity	Database on EU security and military projects	opensecuritydata.eu
Papa Reo	NLP tools and APIs made to serve indigenous communities	https://papareo.nz/#kaitiakitanga
Privacy Rights	Database of online data breaches, meta data and breach information itself (mostly US-centric)	https://privacyrights.org/data-breaches
Turing Institute	Tracked reports of digital rights violations	https://www.turing.ac.uk/news/publications/experiences-online-harms
Wayback Machine	Find historical websites and previous versions of webpages	https://archive.org/web/

DFF Digital
Freedom Fund

www.digitalfreedomfund.org

Postal address
Digital Freedom Fund
Nieuwezijds Voorburgwal 104-108 1012
SG Amsterdam The Netherlands