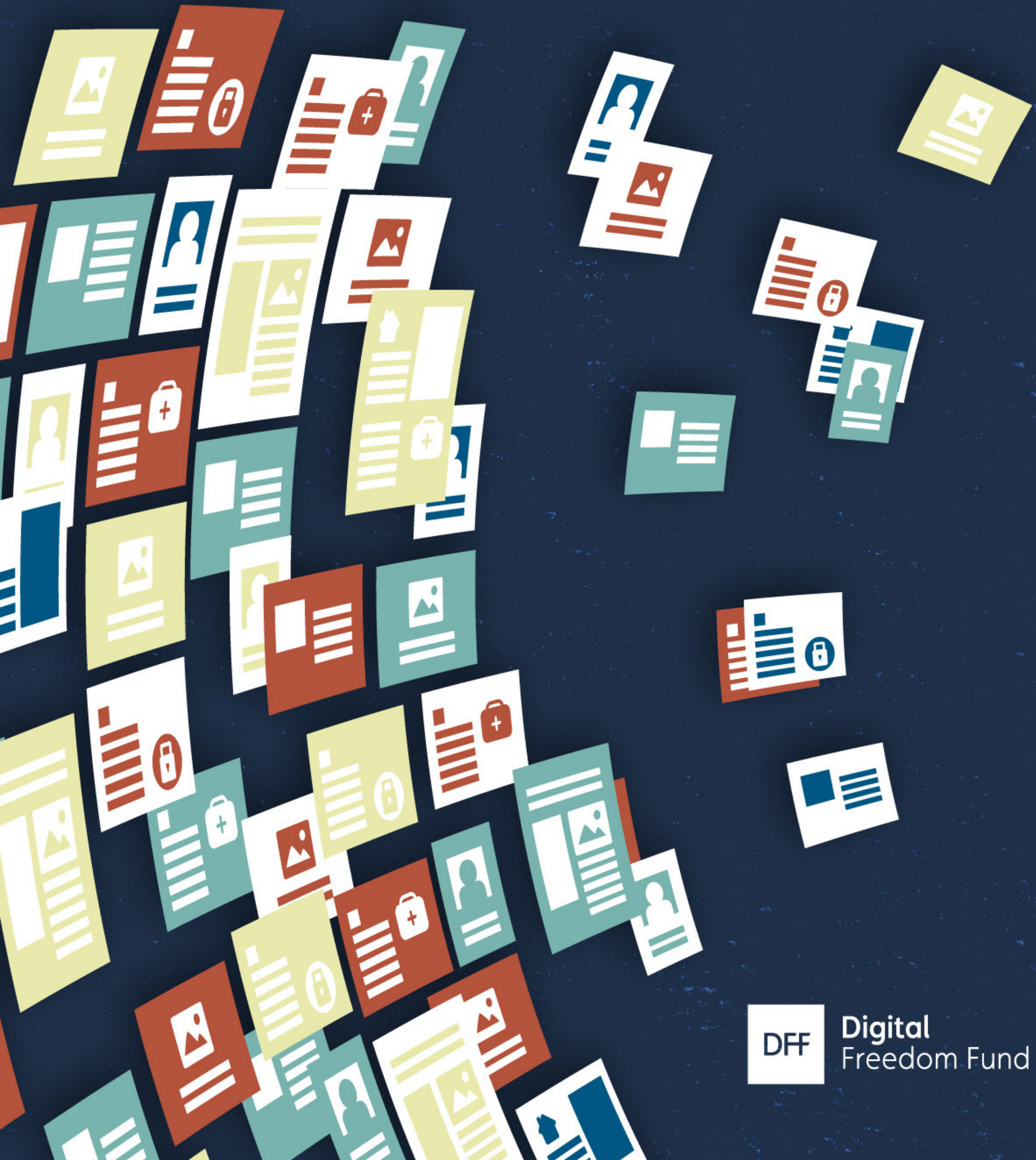
Safeguarding Digital Rights Amidst COVID-19 Through Strategic Litigation on AI



DFF Digital Freedom Fund

Contents

Introduction

As the Coronavirus continues to spread around the world, it has struck a devastating blow to people's livelihoods, exposing vast inequities in health, unemployment, housing, and education, among many other areas. In response, governments are turning to technological solutions to track and monitor populations and enforce safeguarding measures, but this shift precedes the pandemic. Data-driven technologies, particularly automated decision-making systems (ADMs), are increasingly governing key areas of public life, including healthcare and law enforcement. It is in this landscape that governments are introducing COVID-19 technologies in spite of their far-reaching consequences for human rights.

Digital Freedom Fund convened a three-day workshop in order to bring together participants at the forefront of challenging government use of ADMs. The workshop provided a forum for participants across jurisdictions to brainstorm strategies for safeguarding digital rights compromised during the pandemic and for effectively challenging COVID-19 related technologies/ADMs. Specifically, the workshop aimed to:

1. Map out the landscape of technologies adopted both in and outside Europe as a response to the pandemic
2. Draw lessons learned from successful cases against government uses of ADMs
3. Identify next steps and share best practices for challenging COVID-19 technologies through litigation and advocacy

This report summarises key takeaways from this discussion. The first section situates the development of COVID-19 technologies in the context of growing digitisation and automation in how governments interact with the public. Following this, the second section provides a survey of the wide array of COVID-19 apps identified by workshop participants, and the risks they pose to public life and human rights. The third section draws from participants' own experiences of challenging ADMs to offer a toolkit for the strategic fight against unlawful and harmful uses of such systems. On the final day, participants identified several topics as pressing areas for intervention: convincing the courts, preventing the re-purposing of COVID tech, solving tech inequalities, and tackling human and machine bias. The final section summarises perspectives shared and best practices surfaced on how to tackle these challenges.

Section 1: Automation of public life before the Coronavirus

Long before the Coronavirus, governments have been turning to data-driven and automated decision-making systems (ADMs) to execute public policy in key areas, such as healthcare, welfare, and law enforcement. To public administrators and authorities, automation of essential government services offers a cost-effective solution amidst rising government spending and draconian budget cuts to social services. Austerity policies' ripple effects can be seen in the streets, prisons, and at the borders where law enforcement deploys intrusive surveillance technologies to monitor and target marginalised populations and normalise punitive sanctions.

It is in this political climate that governments are turning to technological solutions to the Coronavirus crisis. Workshop participants discussed the importance of situating this approach in the broader context of increasing automation of public life. This section draws from participants' past and ongoing cases against ADMs in healthcare, welfare, migration, policing, and the criminal justice system to outline key issue areas: (1) rise in surveillance; (2) algorithmic bias; (3) repercussions of error rates; (4) failure to carry out impact assessments; (5) lack of transparency and oversight; and (6) disparate impacts.

What does government use of ADMs look like?

The justification given for automating essential public services is that data-driven systems will increase efficiency, save costs, and reduce human bias and error. To that end, algorithmic systems are used to assist decision-making, such as rationing healthcare and assessing defendants' readiness for parole, and predict human behaviour, like anticipating migration patterns and preventing welfare fraud. While these technologies have been touted as "hi-tech" and "AI-powered," many ADMs deployed in government contexts are built on simple statistical models, including decision trees and linear regression.

From data collection to deployment, ADMs involve multiple actors that each have different thresholds and expectations for public transparency and accountability. Though some are created in-house, algorithmic systems used by governments are largely developed by third-party commercial vendors. Throughout their lifecycle, they involve a loose network of actors, including consultants and data scientists responsible for collecting training data sets, conducting evaluations, assessing impact, and carrying out audits. The elusiveness of these networks makes it difficult to pin down who is responsible for design decisions and hold the government accountable for their use of ADMs.

Rise in surveillance through personal and sensitive data collection

As technologies of surveillance become more refined, so, too, do the scope and reach of mass surveillance. ADMs used by law enforcement to target, monitor, and predict populations and human behaviours are becoming more and more intrusive as they combine biometric data with personal data pooled from various government agencies and social media. Workshop participants lamented how quickly and widely such intrusive instruments are adopted with little oversight and expressed serious concerns about the disparate impact they have on communities that are already overpoliced. At a higher level, participants feared the normalisation of state surveillance and erosion of privacy.

Take, for example, technologies of border control designed to detect and monitor migrant movements. European border agencies are partnering with industry to deploy a range of surveillance technologies, including big data analytics, satellite image recognition, and social media profiling, in order to predict and monitor migrant movements. Feasibility studies conducted internally by vendors show that models are most useful when data is specific: the compounding of various technologies (combining drone technology with a facial recognition system, for example) reflects an effort to collect personal data at a more granular level.[1] Participants anticipated that, in the short term, this increase in surveillance will pressure migrants and refugees to make dangerous, high-risk decisions in order to avoid contact with authorities. In the long run, they feared potential misuse and weaponisation of personal data collected in this way.

Thankfully, courts are beginning to put policing technologies in check. Earlier this year, the UK's Court of Appeal found South Wales Police's use of a live facial recognition system unlawful for breaching privacy rights, data protection laws, and equality legislation. The Court's ruling echoed participants' concern around the lack of oversight and a lawful justification for use. This case was litigated at a time when there was a wave of calls to place a moratorium, if not an outright ban, on unlawful uses of facial recognition systems.

Encoding bias

Algorithmic systems do not exist in a vacuum. Creating an algorithm, or a set of mathematical instructions, requires designers to make specific choices about how to operationalise real-world phenomena, by turning them into variables and terms legible to machines. This process of operationalisation is ripe with bias.

Take, for example, risk assessment models used to inform healthcare provision. According to one *Science* article,[2] the model assigned the same level of risk to sicker Black patients as less sick White patients, meaning that less money would be spent on Black patients even where there was the same (or an even greater) level of need. As the authors noted, it is not the case that the model was assigning risk scores purely on the basis of race. Rather, racial bias emerged from the model's use of health costs as a proxy for health needs.

Similar effects of racial and class bias are present in risk assessment tools used for welfare provision. In the Netherlands, the System Risk Indication (SyRI) was introduced to predict the likelihood of committing welfare or tax fraud. While SyRI did not contain race as a variable, its use of neighbourhood data operated as a proxy for race and class that led to a discriminatory impact on poor and migrant communities.

[1] Black C. (2020). *Monitoring Being Pitched to Fight COVID-19 Was Tested on Refugees*, The Bureau of Investigative Journalism: https://www.thebureauinvestigates.com/stories/2020-04-28/monitoring-being-pitched-to-fight-covid-19-was-first-tested-on-refugees.

[2] Obermeyer Z. & Others (2019), *Dissecting racial bias in an algorithm used to manage the health of populations*, Science Magazine: https://science.sciencemag.org/content/366/6464/447.

Repercussions of error rates

Part of what makes algorithmic decision-making systems attractive to governments is their promise of reducing human bias, but their accuracy is widely contested. In particular, the repercussions of error rates pose a serious threat to individuals' safety, freedom, and privacy. Consider false positives in the context of law enforcement: false positives incorrectly assign high risk to certain individuals even though they may be unlikely to reoffend. This means that a recidivism scoring instrument used by judges may deny a defendant's parole request because it falsely assigned her a high-risk score.

On the flip side, incorrectly assigning a low risk score, or false negatives, carries the risk of underserving individuals with need. In healthcare and welfare contexts where risk scoring instruments are used to ration benefits and care, undercalculating individuals' needs can deprive them of the essential services they depend on.

Assessing impact

Given how ADMs' accuracy and reliability are questionable at best, the lack of oversight in creating, deploying, using and maintaining them in contexts of government use is deeply troubling. Participants observed how the "rush to deploy" trumps the standards and protocols that are already in place for justifying, tendering, procuring, and auditing third-party products and services for government use.

In particular, participants identified impact assessments as a key point of intervention for ADMs oversight. It is already standard practice for public administrators to conduct internal assessments for articulating purpose, anticipating harms, and establishing safeguarding practices before executing public policies. ADMs should be subjected to the same procedures in order to ensure that systems are not only lawful, but actually fit for use. Despite their importance, participants' efforts to investigate the presence and scope of impact assessments often ran up against a dead end as governments and courts denied information requests on grounds of trade secrecy and other confidentiality and intellectual property protections.

The tides on government oversight of ADMs are, hopefully, beginning to change. In Liberty's case, the court ruled law enforcement use of live facial recognition unlawful. The ruling found the lack of an investigation by police into potential discrimination based on race and sex was a breach of equality legislation.

In addition to governments, vendors should be held to a higher standard when developing systems for public use. Feasibility studies, in fact, are already an industry practice, but it is an insufficient measure of impact for systems funded with public money for public use. Evaluations of risks and harms, as participants noted, must precede deployment.

Lack of transparency

Part of what makes algorithmic accountability difficult is the lack of transparency. Governments claim that the process of making and designing automated instruments are not public knowledge because they are often created by third-party commercial entities. This trade secrecy protection allows vendors to withhold how and why systems are designed the way they are, which is integral to challenging their function and impact. Participants repeatedly stated how proprietary protections are infringing on the public's right to know.

Proprietary protection also means that governments conceal their tendering and procurement process from the public. Consider iBorderCtrl, the European Commission's research project for border control. Funded with 4.5 million EUR, this system would require those traveling to Europe to take a photo through their webcam or smartphone; the system would then use this photo to extrapolate biometric data to determine whether those travelling to Europe are telling the truth about the purpose of their journey.[3] Human rights advocates filed Freedom of Information (FOI) requests to investigate iBorderCtrl's design and tendering process, but border agencies denied the requests on grounds of commercial confidentiality.[4]

Participants repeatedly pointed at proprietary protection as one of the biggest hurdles to algorithmic accountability. It also means that public administrators responsible for purchasing and implementing algorithmic systems "use them speculatively" without fully understanding their far-reaching implications in public life. The blackboxing[5] of algorithmic systems from, not only the public, but also responsible authorities creates a worrisome public sector trust and dependency on private actors who develop these proprietary systems.

[3] Gallagher R. and Jona L. (2019), *We Tested Europe's New Lie Detector for Travelers — and Immediately Triggered a False Positive*, The Intercept: https://theintercept.com/2019/07/26/europe-border-control-ai-lie-detector/.

[4] Breyer P. (2019), *EU Court to examine secrecy of lie-detector "artificial intelligence" technology:* https://www.patrick-breyer.de/?p=590076&lang=en

[5] Pasquale, F. (2015). *The black box society*, Harvard University Press: https://www.hup.harvard.edu/catalog.php?isbn=9780674970847.

Disparate impacts

The effects of ADMs are far from equal; instead, they reproduce and exacerbate existing structures of inequality.

Marginalised and vulnerable populations are disproportionately targeted and underserved by intrusive and harmful programmes. For example, refugees and migrants were the first to be subjected to European border agencies' invasive surveillance technologies.[6] In the US, automation of healthcare provision for low-income adults with disabilities led many to experience horrific living conditions because the system drastically reduced their care hours.[7]

Data-driven approaches to social problems also dehumanise the people involved. Consider the risk scoring instruments used to determine readiness for parole in the US. Investigation into one widely used instrument revealed how risk calculation relied heavily on one vaguely-phrased question: it asked the prison counsellor to assess the defendant's "disciplinary issues" in a binary response ("yes" or "no").[8] This one question could determine the defendant's release from prison. Codifying readiness for parole into one question, participants remarked, illustrates how the risk calculation fundamentally misapprehends the parole process and dehumanises defendants.

"Like the Hydra's head"

Participants are beginning to see a pattern. Local authorities partner with vendors to repurpose their software and implement them—however flawed, harmful, and unlawful they may be.

Echoing similar observations in their jurisdictions, participants likened the situation to "the Hydra's head. You cut one off, they keep coming." This was the case with SyRI: earlier this year, the District Court of the Hague's ruling found SyRI unlawful, marking an important moment for rights-based resistance against ADMs. Within months, however, the Dutch government introduced a new proposal permitting a similarly, if not even more, intrusive and discriminatory system.

[6] Black C. (2020), EU agencies tested monitoring data on refugees, EU Observer: https://euobserver.com/coronavirus/148185

[7] Eubanks, V. (2017). *Automating inequality: How high-tech tools profile, police, and punish the poor,* New York, NY: https://blackwells.co.uk/bookshop/product/Automating-Inequality-by-Virginia-Eu-banks-author/9781250074317.

[8] Wexler R. (2017), *Code of Silence: How private companies hide flaws in the software that governments use to decide who goes to prison and who gets out,* Washington Monthly: https://washingtonmonthly.com/magazine/junejulyaugust-2017/code-of-silence/.

Section 2: Tech responses to the COVID-19 pandemic: trends and impact

As the previous section made clear, state actors increasingly rely on data-driven systems and predictive technologies to deliver essential services. This trend has accelerated since the outbreak of COVID-19. Faced with a crisis of unprecedented scale and severity, governments across the globe are turning to technology to tackle the virus and facilitate a ‘return to normal’. Over the last six months, this has led to an explosion of new digital tools—from symptom trackers to contact-tracing applications—and the repurposing of existing technology in service of population health management.

The following section presents a high-level typology of some of these tools (see Table 1), before highlighting the common risks they pose to fundamental rights, and how these might be challenged through strategic litigation. Both in and outside Europe, participants noted that widespread technology use was often touted as the only viable exit strategy from COVID-19 lockdowns. If left unchecked, however, many fear that such ‘tech solutionist’ measures erode individuals’ freedoms and compound existing inequalities.

Automation and healthcare during COVID-19

Analytics and data-driven technologies increasingly inform how providers process patient information and deliver care. Even before the pandemic struck, this shift raised serious concerns about data integrity and equitable access to treatment. Public health advocates have long sounded the alarm about ADMs disenfranchising and excluding vulnerable and marginalised populations from accessing quality care. For participants monitoring the use of these tools in healthcare practices, the additional resource strain caused by the COVID-19 crisis has only worsened these discriminatory dynamics.

In the US, algorithmic systems continue to mediate how healthcare professionals ration care and allocate funding amidst resource and capacity strains on life-saving treatments like ventilators.[9] Since the onset of COVID-19, some states have developed “crisis standard of care” guidance to help medical staff determine who should access life-saving treatment. By relying on decision-making tools and policies that prioritise long-term life expectancy, however, such measures were identified as potentially illegally excluding people with disabilities and were changed after administrative complaints.[10]

Similarly, algorithmic tools used to place COVID-19 testing sites reproduce racial and class inequalities that structure unequal access to medical care. Journalists and advocates investigating the distribution of testing sites found that they were being built in historically white neighbourhoods,[11] because data analysis was prioritising population size to determine location, rather than other considerations, like infection risk level and testing need. The use of algorithmic tools by government agencies to distribute COVID-19 relief funds, finally, is also more likely to exclude poor communities of colour.[12] This racial bias is, as the previous section discussed, in great part due to vendors’ use of healthcare costs as a proxy variable for healthcare needs.[13]

Reflecting on these developments, participants underscored how value judgments about who deserves to receive medical care — and ultimately whose life matters— is encoded into the very data systems that distribute it. Looking forward, many anticipated that similar value judgments might inform vaccine allocation if and when they become available, with far-reaching consequences for population health.

[9] Park A. and Kluger J. (2020), *The Coronavirus Pandemic Is Forcing U.S. Doctors to Ration Care for All Patients*, TIME: https://time.com/5825145/coronavirus-rationing-health-care/.

[10] Center for Public Representation (2020), *COVID-19 Medical Rationing and Facility Visitation Policies:* https://www.centerforpublicrep.org/covid-19-medical-rationing/.

[11] Farmer B. (2020), *The Coronavirus Doesn’t Discriminate, But U.S. Health Care Showing Familiar Biases*, National Public Radio: https://www.npr.org/sections/health-shots/2020/04/02/825730141/the-coronavirus-doesnt-discriminate-but-u-s-health-care-showing-familiar-biases.

[12] Bass D. and Tozzi J. (2020), *U.S. Covid Funding Flaw Shortchanges Hospitals in Black Communities*, Bloomberg: https://www.bloomberg.com/news/articles/2020-09-10/u-s-covid-funding-flaw-shortchanges-hospitals-in-black-communities.

[13] Obermeyer Z. & Others (2019), *Dissecting racial bias in an algorithm used to manage the health of populations*, Science Magazine: https://science.sciencemag.org/content/366/6464/447.

Classifying COVID-19 digital tools

Curbing the spread of the virus raises several operational challenges for governments. Beyond providing accurate safeguarding information to the public and effectively tracing infected individuals, public health officials must also incentivise broad behavioural changes in the population. This includes ensuring compliance with social distancing and quarantine rules to prevent further transmission. Spurred on by government calls to action,[14] the onset of the pandemic saw developers in the public and private sectors rush to devise technological solutions to these challenges. These applications—many of which build on existing smartphone infrastructures and technologies of surveillance such as GPS and facial recognition—roughly fall into three categories.

TABLE 1. COVID-19 related applications monitored by workshop participants

Type of tech	What is it for	What it does	Data collected	Responsible authority	Jurisdictions of deployment
Contact tracing	Trace and alert individuals who have come into contact with or might have infected others	Collect location/association data; give exposure notification to end-user; provide risk assessment	Location data via GPS/Bluetooth	Health authority or equivalent Developed in-house or by third party	EU, India, Latin America
Symptom tracking	Help individuals check their symptoms and direct them to appropriate resources	Provide end-user with a diagnosis based on self-reported symptoms; evaluate infection risk; provide guidance	Location data Biometric information Cookies (including browser data, Google analytics)	Health authority or equivalent	EU, Latin America
Quarantine and distance enforcement	Enforce social distancing and home quarantine measures	Monitor users' location and distance from others; report to law enforcement	Location data End-user biometric data (including from photographs)	Law enforcement	Poland, Latin America

[14] Office of Science and Technology Policy, *Call to Action to the Tech Community on New Machine Readable COVID-19 Dataset*: https://www.whitehouse.gov/briefings-statements/call-action-tech-community-new-machine-readable-covid-19-dataset/

Symptom tracking apps

Symptoms checking apps are mobile applications that allow users to check whether they have been infected with COVID-19 based on self-reported symptoms. Some of these applications only collect biographical information related to a patient's age, weight, gender, current symptoms and any chronic or underlying conditions. In other cases, however, the apps require identification and may collect information about users' social background, such as their travel history and contact with other patients. Based on users' answers, most apps then produce a simple diagnosis. Some may also issue risk assessments for at-risk groups based on statistical probabilities, and offer guidance on what actions users should take if infected (e.g. how to get tested and whether to seek further medical assistance).

Symptom checking apps were rolled out early on in the pandemic, particularly as rising cases placed serious strain on medical staff and resources. The Spanish government, for example, touted symptom checkers as a good way to take pressure off hospital hotlines and provide public authorities with an overview of the number of patients requiring medical assistance. The scientific and medical communities also welcomed symptom trackers for its collection of valuable longitudinal data. In places like Greece where COVID-19 tests are expensive or otherwise inaccessible, symptom trackers also offer an attractive alternative to testing. As one participant remarked, while self-monitoring may create an "illusion of control", the diagnoses and risks assessments generated by these applications are based on rudimentary data and risk models and thus far from rigorous enough to ensure public safety.

Contact tracing apps

Contact tracing apps are mobile applications designed to alert users who have been in close proximity with others who have tested positive for COVID-19. The aim of contact tracing is to interrupt virus transmission by ensuring that individuals who come into contact with infected individuals are notified and take steps to contain further spread. While manual contact tracing can be laborious, costly, and time consuming,[15] smartphone apps promise to cut costs and expedite the process by notifying at-risk individuals in early stages of exposure. Several countries across Europe and in Latin America, including Peru, Colombia and Dominican Republic[16] have placed mobile contact-tracing at the heart of their pandemic response strategy.

Beginning in Spring 2020, technologists around the globe rushed to develop protocols for contact tracing apps that could be both privacy-friendly and efficient from a public health perspective. As several workshop participants noted, this sparked an intense debate over the advantages and drawbacks of centralised versus decentralised approaches to data storage. Most contact-tracing apps infer the closeness and duration

[15] Soltani A and Others (2020), *Contact-tracing apps are not a solution to the COVID-19 crisis*, Brookings: https://www.brookings.edu/techstream/inaccurate-and-insecure-why-contact-tracing-apps-could-be-a-disaster/.

[16] Arroyo V. (2020), *Tecnologías de vigilancia para controlar el COVID-19 en América Latina*, Access Now: https://www.accessnow.org/tecnologias-de-vigilancia-para-controlar-el-covid-19-en-america-latina/.

of contact between individuals based on Bluetooth exchanges between mobile devices. Other apps, such as India's Aarogya Setu, go further and trace users' whereabouts using both GPS and Bluetooth data, which are then stored on a government-controlled server. Centralised architectures give health authorities sole control of users' social graph and proximity data, while decentralised infrastructures only store that information on users' devices.

In Europe, Germany initially supported the EU-wide effort to develop a centralised standard.[17] This approach was, in fact, framed as a matter of "digital sovereignty"[18] by French officials who argued that the responsibility to define what constitutes 'contact'; and protect public freedoms should ultimately fall to the state, and not private entities—a "hill that they were ready to die on", according to one participant.

Under pressure from privacy advocates, however, many European countries, including Germany and the UK, eventually adopted the decentralised structure. Growing support from industry actors like Apple and Google also played a crucial role in the process, and other European countries followed suit. In June 2020, for instance, the Norwegian Data Protection Authority decided to impose a ban on centralised processing of personal data collected by the Institute of Public Health's contact tracing app,[19] setting an important precedent for other DPAs. Currently, the Decentralized Privacy-Preserving Proximity Tracing (DP-3T) protocol based on the Google/Apple exposure notification API has become the standard across Europe, with France the only country considered during the workshop that is pursuing its own centralised system, StopCovid.

Quarantine enforcement apps

Finally, quarantine enforcement apps are smartphone applications used by law enforcement to enforce social distancing rules and quarantine measures. Most of them track users' movements based on location data provided via Bluetooth or GPS. In Mexico, Brazil, Chile and Colombia, these data are provided by telecommunication companies and police drone footages. In other jurisdictions like Poland where the applications' uptake is mandatory, the apps enable geolocation tracking by default and automatically share the data with the police. The Polish app also sends users unscheduled requests to upload a photograph of themselves within 20 minutes of receipt to enforce mandatory quarantine. Verification is then established through a mix of facial recognition and GPS geofencing,[20] and individuals who fail to comply face a EUR 7,000 fine and a charge of misdemeanour.

[17] Becker A. (2020), *EU push for coronavirus contact tracing suffers setback*, D Deutsche Welle: https://www.dw.com/en/eu-push-for-coronavirus-contact-tracing-suffers-setback/a-53219372.

[18] Fabre M. (2020), *Souveraineté numérique: le gouvernement se passe d'Apple et Google pour son application Stop Covid, Novethic*: https://www.novethic.fr/actualite/numerique/donnees-personnelles/isr-rse/le-covid-19-fait-ressurgir-l-importance-d-une-souverainete-numerique-francaise-148506.html.

[19] Norwegian Institute of Public Health (2020), *NIPH stops collection of personal data in Smittestopp*: https://www.fhi.no/en/news/2020/niph-stops-collection-of-personal-data-in-smittestopp/.

[20] Ministry of Foreign Affairs Republic of Poland (2020), *Home Quarantine Monitoring by TakeTask*: https://www.gov.pl/web/diplomacy/home-quarantine-monitoring-by-taketask.

Impact of COVID-19 technologies on public life

COVID-19 applications play a vital role in spreading public health messages, but the rush to expand their use without sufficient oversight can have far-reaching consequences for public life. This section summarises these concerns by locating them in the lifecycle of COVID-19 related technologies: from design and ideation, to their various uses and implementation, and long-term consequences beyond the pandemic.

Design and procurement

COVID-19 related technologies, like many ADMs, are increasingly procured from private entities and built on proprietary software. Outsourcing the development and implementation of these apps to third parties, however, raises fundamental questions about whose values and objectives are encoded into these technologies. As public health officials lack a clear understanding of these system's inner workings and their potential for harmful effects, they may not adequately safeguard the public against the various risks they might pose: from biased representations to infringements of privacy.

Another problematic aspect of procuring data-driven systems from third-party vendors concerns responsibility, especially how data from these systems is handled, by whom and for how long. Participants feared that systems built by private entities make it more difficult "to fully understand how private information is being used, shared and disseminated." Whilst some technology vendors are required to perform risk assessments to determine whether they comply with principles of necessity and fairness, for example, opaque agreement structures obscure these processes.

Beyond that, public-private partnerships behind closed doors also compromise the public's right to know. Consider, for example, what happened in Greece: the public discovered the roll-out of a smart policing programme from the vendor's press announcement about signing a contract with the Hellenic police while the Greek authorities kept it a secret. An added complication arises with the multiplication of private vendors. In Argentina, one participant highlighted, different provinces steered away from the government strategy and started launching their own apps "without any alignment with what the national health ministry was pursuing."

Implementation and use

Despite their widespread deployment, there is little consensus about the accuracy and efficacy of self-monitoring and contact tracing apps. Research shows that the impact of tracking apps in reducing the number of infected individuals is more limited than random testing.[21] Likewise, experts suggest that tracking apps need an uptake of at least 60% of the population to meaningfully slow down virus transmission[22]—a far cry from current adoption levels across the world.

Besides, contact tracing apps only act as imperfect proxies for exposure as they wrongly assume that all users carry the same risk of transmission. Mobile applications are prone to failure and "buggy functionality" risks producing false positives by tracing individuals who do not present a high risk of infection. This is particularly alarming in jurisdictions where a positive diagnosis is conducive to mandatory home quarantine or where smartphone applications serve as 'immunity passports' for access to work sites and public transportation. Such technical shortcomings can also put unnecessary strain on health care services by sending individuals to hospitals who do not need care.

Even if they were effective, COVID-19 apps exacerbate the digital divide by excluding people who do not have access to smartphones or strong connectivity. In regions where smartphone penetration is high, the main barrier to adoption is often public trust in the technology itself. However, in large parts of the world, smartphone ownership is far from a given, due to cost or lack of need. A jarring example of this is India, where less than half of the population owns a smartphone or has internet access, despite the government's contact tracing app being essentially mandatory.[23] And even in regions with good mobile penetration rates, as is the case across most of Latin America, connectivity issues are still common.

Finally, people's devices may simply not be compatible with COVID-19 apps. Some governments offer exemptions for those who do not own a smartphone, under strict conditions: in Poland, individuals must submit an official declaration to the authorities, with non-compliance considered a criminal offence punishable by 3 years in prison. Participants remarked that by equating individuals with smartphone devices, governments ultimately shape who is "hypervisibilised" and who is "rendered invisible" by their policy decisions. And several of them likened the lack of attention paid to these disparities to a "failure of government accountability and due process" for those left out.

[21] European Parliament (2020), *Parliamentary Questions: Tracking and tracing citizen movements in response to the coronavirus,* Question for written answer E-002473/2020: https://www.europarl.europa.eu/doceo/document/E-9-2020-002473_EN.html.

[22] University of Oxford (2020), *Digital contact tracing can slow or even stop coronavirus transmission and ease us out of lockdown:* https://www.research.ox.ac.uk/Article/2020-04-16-digital-contact-tracing-can-slow-or-even-stop-coronavirus-transmission-and-ease-us-out-of-lockdown.

[23] Christopher N. (2020), *India made its contact tracing app mandatory. Now people are angry,* Wired: https://www.wired.co.uk/article/india-contact-tracing-app-mandatory-arogya-setu.

Beyond creating barriers to care, COVID-19 apps also run the risk of normalising intrusive data collection practices. Several of the symptom tracking apps discussed fail to secure informed consent from users and show little concern for the GDPR principle of data minimisation. Extreme examples of this are the Spanish and Greek self-monitoring apps, which require access to features and data that far exceed what is required for the purpose of triage, such as microphone access and internet browsing data.

Participants feared that forcing such digital solutions onto the public impinges on their human rights. Contact tracing app use remains ‘voluntary’ in most jurisdictions, but the pressure governments place on the public, at times with punitive sanctions, leave individuals with no real choice. In some parts of India, for example, not having the national contact tracing app downloaded limits access to basic services, creating a false incentive to adopt the technology under an illusory notion of ‘choice’. And while that decision has now been challenged in court, for the better part of the pandemic (so far) uptake was mandatory for people living in ‘quarantine zones’ else they faced prison sanctions or fines.

Beyond the pandemic

Looking ahead, participants anticipated several long-term repercussions of privacy intrusions normalised during the pandemic. First is the repurposing of data gathered by COVID-19 technologies. Most of the technologies discussed thus far have been deployed under opaque “state of emergency” provisions. Under these circumstances, what will happen to the information collected through these apps is unclear.

For many, relationships between industry and law enforcement agencies in particular create strong data sharing concerns. If those relations fall outside the GDPR’s territorial scope, there may not be clear limits or binding policies on the lifespan of government and provider databases. But even in countries with a strong data protection infrastructure and vocal watchdogs, participants worried that authorities could be vague about which agencies or third-party vendors would have access to that data.

Participants also anticipated that COVID-19 technologies would, through gradual shifts in objectives, facilitate broader surveillance of society at large. They are already witnessing normalisation of thermal cameras and face mask monitoring in public spaces and the expanded use of technology in school to monitor students during exams, for instance. In countries with a history of authoritarianism, expanding the channels of surveillance is particularly concerning.

Reflecting on these developments, participants forecasted that COVID-19 surveillance technologies may have deleterious consequences on users’ mental health in the long-run. By requiring users to “constantly” interact with their devices to check into places, monitor personal health, and share data with law enforcement, many COVID-19 apps may indeed cause considerable stress and feelings of being under sustained observation.

Section 3: Litigation Tactics & Tools

Drawing from workshop participants' own hard-earned lessons of combating harmful uses of ADMs in public life, this section offers some tips and recommendations to consider when taking strategic litigation on these issues. The strategies outlined below range from utilising procedural and administrative legal tools, such as FOI and General Data Protection Regulation (GDPR) data subject requests, to rhetorical and framing choices, like foregrounding the human experience of ADMs to educate judges on the implications and potential harms of sociotechnical systems. Participants acknowledged that many of these tactics may be sidelined or halted as a result of the budget, staff, and resource constraints placed on litigators during the Coronavirus pandemic.

Freedom of Information (FOI) requests

Freedom of information (FOI) laws allow the general public to access data held by national governments and their agencies. Members of the public can file FOI requests to specific public authorities to access information held by them about ADMs in use, including how they function, the government and commercial actors responsible for deployment, and their impact on public life. In jurisdictions with a strong foundation for public transparency, FOI requests can be an incredibly valuable evidence gathering technique.

FOI requests, for example, are a useful way to investigate the presence and scope of impact assessments. Under the GDPR, governments and vendors developing projects that pose a high risk to people's personal information are mandated to conduct a data protection impact assessment (DPIA). This means that apps that involve extensive processing of personal data, like COVID-apps, or large-scale monitoring of public areas, like facial recognition systems,[24] should theoretically have been subject to a DPIA prior to deployment. Failure to produce evidence of having conducted such an assessment, following a FOI request for the DPIA, can be a violation of the law. In this case, advocates and citizens can file a formal complaint to a data protection authority (DPA).

Trade secrecy protection is the biggest hurdle to using FOI requests. Participants expressed frustration at how easily and loosely exemptions were applied on grounds of proprietary protection, including, in some cases, information as basic as vendor name and instruction manuals. Such exemptions, frustrating as they may be, can be instructive. First, it reflects how government agencies and courts primarily see ADMs as a piece of technology, rather than sociotechnical systems that shape public life. This relegates data-driven systems as mere "tools" purchased from commercial entities, and secures them proprietary protection. It also demonstrates how ADMs are blackboxed even from those responsible for implementing them, and captures the knowledge imbalance between government users and third-party vendors.

Proprietary exemptions propelled participants to get creative about gathering evidence. Participants encouraged each other to interpret "information" very broadly to extend beyond technical information, like training data and source code. Often, they found useful evidence emerging from unlikely sources, such as vendors' promotional materials and responsible government agencies' public relations (PR) strategies. This approach also played a rhetorical role: by obtaining information from "non-technical" sources, participants were able to emphasise the social impacts of ADMs.

While "information" can be broadened, participants underscored the importance of striking the right balance between "too specific" and "not specific enough." Some participants have had success when their requests included specific names and other contact information of government actors, contract provisions, and drafts of documents used to obtain permission for implementation. Others cautioned against being too specific as it may disincentivise a response.

[24] European Commission, *When is a Data Protection Impact Assessment (DPIA) required?* https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/when-data-protection-impact-assessment-dpia-required_en.

To make the most out of FOI requests, participants underscored the importance of understanding the information sharing ecosystem of key stakeholders and jurisdictions. ADMs often involve multiple agencies to move from training data collection to deployment, and each agency has different thresholds and expectations for public transparency. Understanding the relationship between responsible agencies, including how information flows internally and externally across departments and chains of custody can inform which agencies to target in filing FOI requests. For example, a FOI request about an algorithmic medical instrument may be more successful when filed to the medical examiner, rather than the police. Relatedly, understanding responsible agencies' procurement and oversight processes can be helpful. "Following the audit trail" to identify what kind of information to obtain, as one participant explained, can be fruitful.

Another useful tactic is establishing and maintaining rapport with agency personnel. FOI requests are filed to government agencies, but decisions are made by responsible individuals. One participant who filed several FOI requests to the police department over the years found that the existing relationship with individuals at the department played a part in the requests' success.

FOI requests filed by members of the public can sometimes lead to different responses. One participant worked together with her clients to file a FOI request: she found the response her client received to be much more substantive than hers, and used this information to build her case.

GDPR data subject requests

The GDPR aims to give individuals protection and control over their personal data and how it is used, while holding entities and organisations responsible for processing them ("data controllers" and "data processors") accountable to standards of transparency and due diligence about their scope of use.[25]

GDPR subject access requests (SARs) offer several advantages over FOI. Since the pandemic began, public administrators are sidelining or simply refusing to fulfil FOI requests at this time.[26] The European Court of Justice, in contrast, is still obligated to uphold data protection legislation, which means that vendors and other stakeholders will be held to the same standard during the pandemic. Unlike FOI, which often returns generic or fragmented evidence, data subject requests can be very specific, allowing participants to build a stronger case to take to data protection authorities. This makes GDPR access requests an effective evidence gathering tactic for individuals and any civil society organisations acting on their behalf.

[25] European Parliament and Council of European Union (2016), *General Data Protection Regulation: Regulation (EU) 2016/679:* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

[26] Salame R. and Zweig N. (2020), *Public Access to Information Suffers Under Coronavirus,* Columbia Journalism Review: https://www.cjr.org/analysis/covid-19-pandemic-foia.php.

SARs are particularly useful to help uncover a technological system's function and use, including authorities responsible for data collection and processing. Under Article 15's 'right to access', any members of the public may request a copy of their personal data from data controllers. Where data is indeed being collected, data subjects can then request additional information on exactly what type of information is being processed about them and for what purpose.

Filing these requests can also be instrumental in building an effective legal challenge, should government agencies or vendors fail to comply with GDPR transparency requirements. Articles 13 and 14 enshrine a data subject's 'right to be informed,'[27] which obligates data controllers to provide individuals with clear and concise information about what they do with their personal data. Outright denial of these requests violates EU data protection laws and can be challenged before the courts.

During this discussion, participants also remarked on how data subject requests can educate the public about their data ownership. Encouraging individuals and human rights advocacy groups to file access requests can elevate the sense that "people have ownership over their data." Requests filed by members of the public could also compel public officials to "take these claims more seriously."

Despite the clear opportunities they present, GDPR access requests present some limitations and risks. An obvious downside to this approach is that participants cannot pursue it when a technology is still at the ideation or developmental stage, as evidence is simply non-existent. If courts believe data controllers to be acting in good faith, another risk is that they "give in too early" and respond to a data protection complaint by simply asking controllers to conduct due diligence. This could restrict litigants' ability to escalate a case to higher courts with human rights violation claims, thus lessening the overall impact of litigation.

Non-legal investigation tips

Sometimes, looking beyond the government can bring useful information. Technologists, whistleblowers, and experts offer valuable insights and skills for discovering and investigating ADMs. Participants found reverse engineering similar systems with technical experts especially useful in cases where not much is known about the specific algorithmic tools in use. Discovery of border control technologies in the UK, for example, emerged from reverse engineering surveillance systems.

Collaborating with grassroots groups, frontline practitioners, and advocacy organisations similarly opens new opportunities for evidence gathering. They are first to witness and experience the human costs of ADMs in public services. Should a discovery

[27] European Parliament and Council of European Union (2016), *General Data Protection Regulation: Regulation (EU) 2016/679, Article 13*: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

lead to a case, partnership with these groups can also contribute towards identifying plaintiffs. Participants who have challenged ADMs in court attributed the strength of their cases to meaningful partnership with civil society and frontline groups.

Other key collaborators are journalists and interdisciplinary researchers across the computational and social sciences. Many are conducting their own investigations into ADMs. Working together with them, rather than starting from scratch, is time- and resource-saving. When an investigation escalates into a case, bringing in experts who can translate technical complexities and centre their analysis on ADMs' impact on people's lives can play a helpful role in educating judges and the broader public.

Monitoring procurement and tendering processes, with special focus on key vendors, is another noteworthy approach to evidence gathering. Participants who have previously challenged ADMs in court discussed how they are starting to identify recurring commercial actors across jurisdictions. Keeping track of these recurring vendors can illuminate their inner-working and internal logics, relationship with government agencies, and terms of contract. Participants also noted how it creates forecasting insight so they can anticipate future contacts and developments and strategise in advance. Finally, it creates opportunities for collaborating with a diverse group of stakeholders.

Public campaigning

Participants discussed the benefits of public campaigning and advocacy as part of a broader strategy of challenging ADMS in public life. Public campaigns aim to raise public awareness and shift perspectives about the social impacts of data-driven technologies, and publicly put pressure on responsible authorities This typically involves collaborating with journalists, grassroots groups or supportive politicians to amplify a shared message and bolster public support for a cause.

Public campaigns often begin with a discovery and investigation phase to effectively communicate how individuals and society at large are affected by a specific programme or technology. Advocates may then opt to wage short-term digital campaigns over social media or engaging in more traditional advocacy through the publication of white papers and lobbying of decision makers.

Public advocacy, several participants pointed out, is an effective way of spotlighting and locating key issues. Concepts like privacy and transparency have legal currency, but compelling stories that articulate how and why automated systems have harmful effects on people's lives are much more powerful in changing public perception. In doing so, successful media campaigns can demystify automated systems and foreground the impact they have on people's lives. Participants cautioned against overreliance on the press, however, especially in jurisdictions under authoritarian regimes where public media acts as a “mouthpiece for the government.”

Striking the right tone is crucial for public campaigns. Where lawsuits are concerned, it is a delicate balancing act between holding governments and vendors accountable in public and avoiding antagonising the opposite party. Take, for example, the case against the use of live facial recognition technology in the UK. Litigants worked closely with local communities and journalists to call for a ban on facial recognition technology. They opted to take a wider policy position than the one they were arguing in court, where they were claiming that facial recognition technology could be used lawfully and in a human rights' compliant way but that the police were simply not using it that way. Liberty's strategy paid off, as they successfully halted the discretionary use of this technology by the police.

As impactful as public campaigns can be, they require extended planning and technical knowledge on the part of advocates. When an individual reaches out with a tip about the discriminatory impact of a system, the key is "building evidence of it being systemic, not just individual." When advocates "do not have data" or sufficient understanding of a technology's inner workings, however, they may struggle to establish its harmful or discriminatory effects. This is often compounded by the difficulty in asserting privacy violations or discrimination as tangible public concerns when the impacts of COVID-19 on human life are far more immediate and material.

Another tangible risk of challenging ADMs in the public eye is gaining unwanted support from political groups. For example, taking a strong position against COVID apps may invite "strange bedfellowship" from far-right extremists or anti-vaccine communities.

Public campaigns also require substantive funding and resources. Civil society organisations are operating under capacity constraints, especially in light of the Coronavirus. Unlike the tech companies building these systems, which have a lot of power and money to lobby on these issues, civil society organisations are beholden to grant funding. Reflecting on these limitations, participants suggested that in some cases, it might be more impactful to lobby stakeholders and decision makers directly.

Building a case

Bringing a case to court is not an easy decision: litigation is costly, time- and resource-consuming, requires strong evidence, and can impose serious strains on the plaintiff. Amidst the ongoing Coronavirus outbreak, many civil society groups that fund and support litigation efforts are operating with additional resource constraints. Participants agreed that building a case must be a decision made wisely and intentionally, and outlined factors to consider.

First and foremost is the question of readiness. Participants insisted that a strong case must be built on sufficient evidence or fact patterns. The tactics mentioned earlier in this section can be instrumental in gathering evidence, but a case in court must be able to demonstrate systematic disenfranchisement of the public's rights.

The question of jurisdiction and time are crucial as well. To maximise the chances of winning in court, cases should be brought up in optimal jurisdictions with strong precedents and sympathetic judges. It is ideal for the case to be the focal point of judicial and public attention. Participants noted how the courts and journalists are being less responsive than usual due to the ongoing pandemic.

If the circumstances of bringing a case are met, the next important factor to consider is the availability of a strong legal team. As participants discussed, litigation that challenges governments are often brought by civil society organisations that are already operating under budget, staff, and resource constraints. Participants advised collaborating with other interest groups, as was the case with a coalition of privacy groups that worked together to challenge SyRI in the Netherlands, or reaching out to private firms that can lend their document reviewing infrastructure.

Participants' wishlist for building a strong case also included having suitable plaintiffs and experts. Suitable plaintiffs who can tell compelling stories that articulate the harmful impacts of ADMs and connect the technologies to broader social issues can have a powerful influence on the course of the suit. However, as participants acknowledged, litigation places serious burden on the plaintiff's wellbeing and privacy. For example, participants who have challenged facial recognition systems in Europe discussed how communities of colour most directly targeted by the surveillance technology were reluctant to bring a case because it could place them under unwanted public attention. Participants emphasised the importance of being mindful of the trade-offs plaintiffs have to make. With Corona apps, participants questioned whether it would even be possible to identify individual plaintiffs, since they pose collective, rather than individual harms.

Similarly, experts who can translate ADMs' complex technicalities in accessible terms to the court are much desired. Experts should also be able to centre their analyses of ADMs on how they impact people's lives to aid litigants in convincing potentially sceptical judges.

When successful, a strong case with compelling stories can have legal, political, and social impacts. Most directly, it can put a stop to harmful ADMs, hold governments to account, and set an important precedent. It can also change judicial authorities' and governments' perception of ADMs by articulating clear harms done to individuals. In a similar vein, a strong case has the potential to change public narratives through media coverage.

Section 4: Deep dive into the challenges ahead

From workshop discussions, participants identified the following topics as the litigation challenges ahead for safeguarding digital rights amidst COVID-19:

1. Convincing the courts
2. Preventing repurposing of COVID tech
3. Solving tech inequalities
4. Tackling machine and human bias

Reflecting on their past experiences, participants shared lessons learned, successful tips, and key takeaways to overcome these drawbacks and challenges.

Convincing the courts

Judges and lawyers often lack a sociotechnical understanding of how ADMs work and impact people's lives. As a result, judges may be sceptical about or "have a blind spot" on the real-world impact of automation, which affects the kind of cases pursued and how those cases are ruled. To help overcome this challenge, participants suggested the following practices and tips for persuading, educating, and engaging judges:

1. Develop sociotechnical curriculum for judicial education and training
2. Use independent experts who can explain the technical aspects in accessible language and foreground human rights perspectives in their analysis
3. Use third-party "explanatories" or interventions that are written in accessible language with clear examples of familiar technologies
4. Be judicious about focusing on technical complexities. Just because technology is involved, it does not mean that foregrounding the technical is necessarily useful from a strategic standpoint

Preventing Re-purposing of COVID Tech

Governments often tout invasive surveillance technology, such as facial recognition, as the most efficient way to uphold public safety and security in times of crisis, a "necessary intrusiveness" under the guise of public good. As we move beyond the pandemic, however, this poses the risk that, if left unchecked, public bodies will normalise mass surveillance by repurposing these technologies post COVID. To safeguard against this, participants suggested the following ideas:

1. Clearly articulate the impact of surveillance technologies to the public through impactful public campaigns
2. Collaborate with privacy organisations across jurisdictions to collectively challenge the narrative that digital tools were a boon to pandemic response
3. Challenge the efficacy of these tools for the purposes they were implemented for
4. Improve the transparency at this stage so that we can find out later whether standards have been upheld
5. Push for more stringent law, policy and guidelines around "emergency exceptions" and the limitations they place on technology use

How to Solve Tech Inequalities

The pandemic has laid bare the extent of systemic injustices running across all sectors of public life: from unequal access to education and healthcare to rampant welfare disparities. Government's over-reliance on technology in these trying times only compounds these inequalities, by excluding large swaths of the population from much-needed resources, safeguarding information and quality care. To mitigate some of these pernicious secondary effects, participants proposed the following strategies:

1. Stop treating affected populations as "afterthoughts" and strive to build their concerns and viewpoints into policy-making decisions upfront
2. Gather systematic evidence about exclusions and harms caused by technology to raise public attention on the issues
3. Push for better data collection and transparency on the part of decision makers around health and economic disparities
4. Elaborate user-friendly guidelines and educational programs to better inform the public and the courts about the implications of developing systems that automate human decision making
5. Engage in meaningful organising to "humanise" or "demystify" algorithms, counterbalance surveillance, and challenge systemic lack of accountability

Tackling machine and human bias

The question at the core of challenging unlawful and harmful uses of ADMs is one of accountability: who, or what, should be deemed responsible and to what degree? The first step to addressing these questions, according to participants, is to understand how biases are built into automated tools to structure their disparate impacts. Discussions about algorithmic bias, however, can easily fixate on the technical aspects (i.e., bias in source code and training data) at the risk of minimising their impact on people and syphoning ADMs from larger structural inequalities at play. To centre on the human side of algorithmic discrimination and focus on the question of responsibility, participants shared the following tips:

1. Expand the list of stakeholders beyond "governments" and "vendors" to better reflect the wide ecosystem of ADMs use
2. Map out the lifecycle of systems in use (i.e., ideation, testing and design, implementation, use, aftermath) to demystify them, expose underlying assumptions, and anticipate harm

3. Develop a basic understanding of how operationalising complex social phenomena into algorithms reproduces biases and stereotypes
4. Frame analysis of ADMs around compliance with human rights principles and standards
5. Disaggregate the notion of responsibility so that the different stakeholders involved in ADMs implementation can be held to proportionate levels of responsibility
6. Position the fight against ADMs in concert with other efforts for protecting privacy and public transparency

About the Digital Freedom Fund

The **Digital Freedom Fund** supports strategic litigation to advance digital rights in Europe. With a view to enabling people to exercise their human rights in digital and networked spaces, **DFF** provides financial support for strategic cases, seeks to catalyse collaboration between digital rights activists, and supports capacity building of digital rights litigators. **DFF** also helps connect litigators with pro bono support for their litigation projects. To read more about DFF's work, visit: **www.digitalfreedomfund.org.**

The project on "Safeguarding Digital Rights Amidst Covid-19 Through Strategic Litigation on AI" was made possible thanks to the support of the Foundation for Democracy and Media. DFF receives organisational support from Open Society Foundations, Adessium Foundation, Luminate, Ford Foundation, Fondation Nicolas Puech, and The Sigrid Rausing Trust.

About the Authors

Kate Sim is a doctoral candidate at the Oxford Internet Institute where she researches the intersection of sexual justice and emerging technologies through ethnographic methods. She is also a Digital Ethnographer for the European Commission's Next Generation Internet project.

Nahema Marchal is a doctoral candidate at the Oxford Internet Institute, University of Oxford and a researcher for the Computational Propaganda Project. Her research examines the relationship between social media and political polarization and the implications of digital technologies for public life.

Democracy & Media Foundation **Stichting Democratie & Media**

DFF Digital Freedom Fund

E-mail:
info@digitalfreedomfund.org

Postal address
Digital Freedom Fund
P.O. Box 58052
1040 HB Amsterdam
The Netherlands