WORLD ECONOMIC FORUM

Pathways to Digital Justice

WHITE PAPER

SEPTEMBER 2021

Images: Getty Images

Contents

Preface

Sheila Warren
Deputy Head, Centre for the Fourth Industrial Revolution; Member of the Executive Committee, World Economic Forum

Evîn Cheikosman
Policy Analyst, Crypto Impact and Sustainability Accelerator (CISA); Manager, Global Future Council on Data Policy, World Economic Forum

Sina Fazelpour
Assistant Professor of Philosophy and Computer Science, Department of Philosophy and Religion and the Khoury College of Computer Sciences, Northeastern University

The World Economic Forum's Global Future Council on Data Policy liaised with the Global Future Council on Media, Entertainment and Sport and the Global Future Council on AI for Humanity, in collaboration with an advisory committee consisting of experts from around the world, to make the case that a new policy framework is needed to effectively address issues of justice that arise in a range of digital contexts. In doing so, the hope is that legal and judicial systems can then evolve to embed redress mechanisms which enable the creation of a data ecosystem that protects individuals and is accountable to them.

In using this white paper to guide policy efforts towards combating data-driven harms, governments across the globe can feel confident that the proposals are fit for purpose for our current digital era, are trauma-informed[1] and victim-focused. We anticipate that this white paper, in conjunction with the complementary tip sheet and resource guide, will serve as instructive, inclusive and beneficial resources for any government addressing the next generation of technology-enabled harms.

It is critical to note that with respect to data-driven and predictive technology, the aim is not to claim that technology itself is the sole source of harm, or that regulating technology is the only viable solution. Technology will continue to advance, and indeed will have the potential to bring positive benefits to society. This reports calls attention to the inadequacy of legal and judicial systems, as well as the quasi-legal and judicial systems that platforms offer, to address the types of harms that arise from these technologies.

The World Economic Forum's network of Global Future Councils is the world's foremost multistakeholder and interdisciplinary knowledge network dedicated to promoting innovative thinking and strategic insights to shape a more resilient, inclusive and sustainable future. The network convenes more than 1,000 of the most relevant and knowledgeable thought leaders from academia, government, international organizations, business and civil society.

A note from the Global Future Council on Data Policy Co-Chairs

JoAnn Stonier
Chief Data Officer,
Mastercard

Marietje Schaake
Director, International Policy, Cyber Policy Center, Stanford University

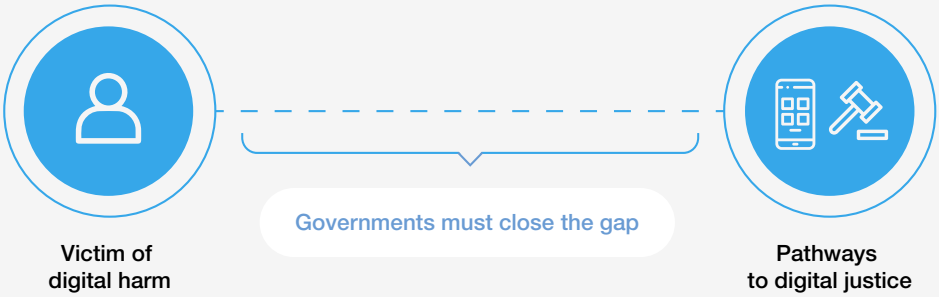
The outbreak of COVID-19 has highlighted the global dependence on digital technologies and networks for economic, health, educational, cultural and scientific endeavours. That reliance has also created new vulnerabilities for data to be weaponized to spread misinformation and disinformation and create societal divides. Data-driven technologies create new injustices and are a growing source of personal and community harm. Meanwhile, accessing pathways to justice and redress in today's digital society is difficult and near impossible in some cases.

Lawmakers in various jurisdictions have tried to address some of the issues of digital harm related to democracy, public health and data privacy.

However, justice and redress in light of new types of harm need better corresponding regulatory and judicial protections. Judicial and regulatory systems also need to evolve to protect the rights of individuals, communities and ecosystems as our data norms and values continue to mature and data ethics are better defined.

There is an accountability gap related to the rise of digital harms, which is enabling a data ecosystem in which bad actors are free to behave with impunity. The regulatory response to such behaviour is insufficient, leading to a secondary problem: that victims of digital harm lack a clear pathway to justice.

FIGURE 1 **We need to acknowledge the accountability gap with regard to digital harm**



The aim of this white paper is to elevate the urgency of creating data ecosystems that reduce harm and that, most importantly, are accountable to people. Governments and other stakeholders across the globe who wish to combat injustices and harms to their citizens are invited to use this white paper to start their work in defining the core elements of effective policies towards digital justice.

As co-chairs, we thank our fellow Global Future Council on Data Policy members, the larger project community and the staff at the World Economic Forum for their contributions to this white paper on Pathways to Digital Justice.
We hope that law- and policy-makers find the data, insights and recommendations helpful, and we look forward to feedback and continued engagement about this important topic.

Executive summary

This paper provides two approaches to creating clear pathways to digital justice that governments can take.

FIGURE 2 **Summary of the overall shifts that need to be made**

What isn't working ⟶ what needs to happen:

1. No recourse for victims ⟶ clear recourse for victims
2. Complete lack of support and awareness ⟶ trauma-informed and victim-centred support
3. "Problem of many hands", online anonymity, no responsibility ⟶ clear accountability
4. Burden is on victims to figure out their own route ⟶ victims have clear steps to navigate justice process
5. Governments and stakeholders are not taking effective, cohesive action ⟶ governments develop new multistakeholder ways to modernize access to justice
6. Treating digital harm as a matter exclusive to privacy ⟶ treating digital harm as a human rights violation that encompasses all forms of personhood and self-determination
7. Fragmented and confusing jurisdictional approach to personal data protection ⟶ modern, flexible data-privacy standard that facilitates international trade while safeguarding human rights

❝ Malicious actors spread information (true as well as false) not merely to negotiate and assert their own societal values but also to target and harm other individuals and groups.

From doing business to staying informed of the news, we rely on access to information 24/7. But what happens when that information targets us in an unjust manner with real-life consequences? We often refer to disinformation as a political phenomenon, but it is becoming increasingly more personalized. Malicious actors spread information (true as well as false) not merely to negotiate and assert their own societal values but also to target and harm other individuals and groups. And the platforms and technologies upon which we increasingly depend often serve to amplify such harms.

The overall objectives of this white paper are to provide governments with: (1) a holistic perspective of the harms that data-driven technologies[2] perpetuate: (2) key failures in global legal and judicial systems with regard to digital justice issues; and (3) recommended pathways to digital justice that lawmakers can develop to better protect individuals and communities.

Part 1 of this white paper explores the myriad of technology-enabled harms, whether they emerge from the misuse of real information or deepfake videos, and how these harms are exacerbated by technology ranging from recommendation algorithms to armies of bots used to manipulate them.

Part 2 contextualizes the notion of digital injustice as a matter of corrective justice, which is a way to attain redress for past actions. In particular, corrective justice is ideal for resolving the types of unpredictable harm that tend to come from data-driven and predictive technologies. For

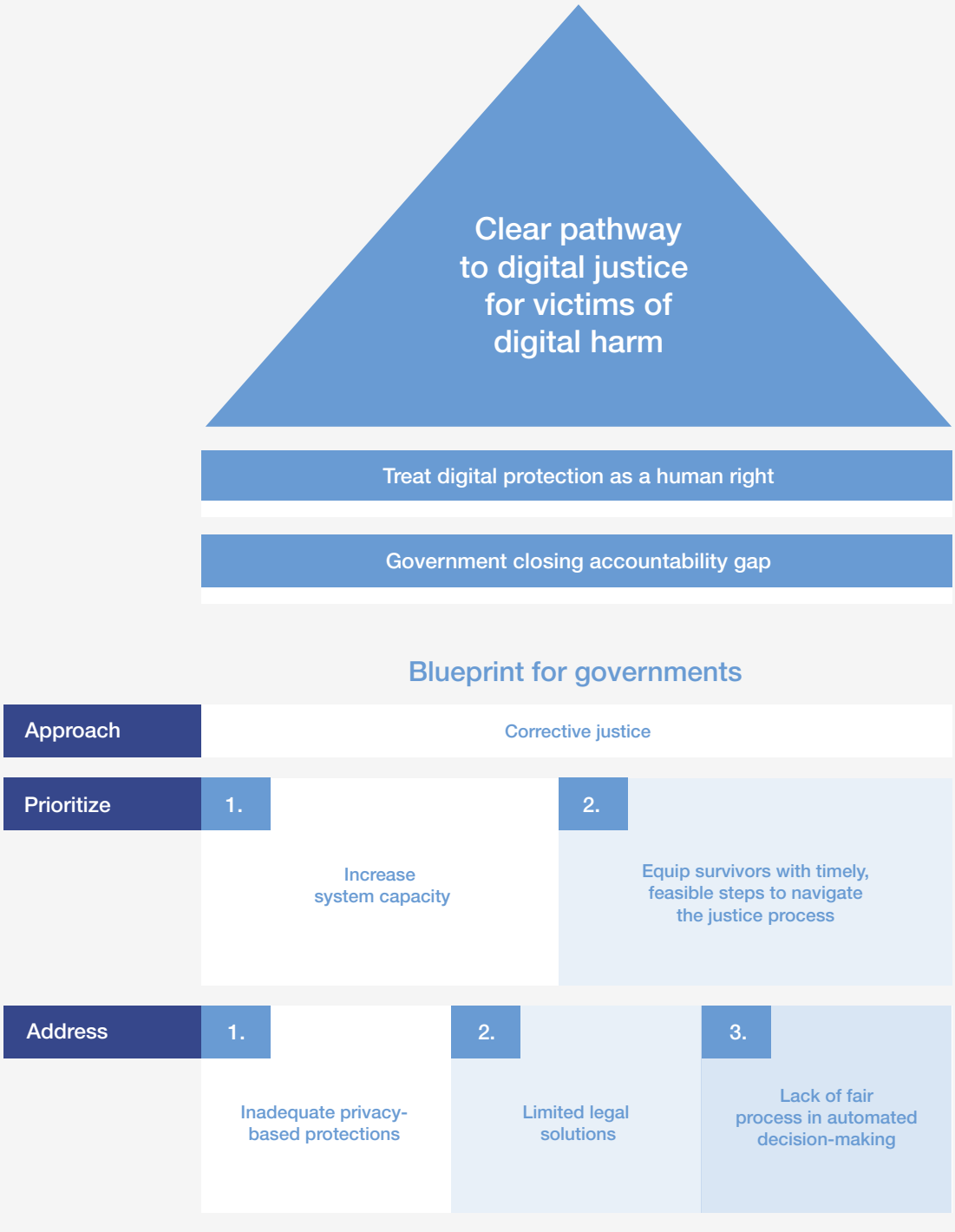
example, corrective injustices occur when we lack the necessary capabilities to detect harms, when there is no accountability or when there are no effective pathways for redress.

Part 3 addresses the three main reasons why justice is hard to achieve. Due to our fragmented and under-resourced legal and judicial systems, and the jurisdictional challenges of regulating international communications platforms, technology-enabled harms have been able to flourish in an environment that defaults to inadequate privacy-based protections, limited legal solutions and a lack of fair process in automated decision-making.

The white paper concludes with two multistakeholder approaches to creating clear pathways to digital justice that governments can take. It underscores the pressing needs to: (1) modernize the judicial system's capacity to adjudicate more claims; and (2) equip survivors with timely, feasible steps to navigate the justice process in the event of some form of harm.

We are proud to present this white paper to you in the spirit of shared progress towards an accessible, trauma-informed and victim-focused pathway to digital justice.

FIGURE 3 **Theory of change**



Part 1

We have a problem

In cases of digital harm, what often seem like violations of privacy are in fact violations of self-determination and personhood.

> We have a problem on our hands as a society. Even though we are increasingly engaging more with technology, the pathways to protect us from the harms that come from those exchanges do not exist.

Ava Rose, TikToker who experienced cyberbullying

Rana Ayyub, investigative journalist and writer who experienced deepfake pornography

Gibi, ASMR YouTuber who experienced deepfakes, online harassment and non-consensual pornography

Robert Julian-Borchak Williams, wrongfully accused by an algorithm of a crime

These are the names of victims[3] of technology-enabled harms who have gone public with their stories. For every victim who has publicly shared their story, there are many who have remained silent.

We have a problem on our hands as a society. Even though we are increasingly engaging more with technology, the pathways to protect us from the harms that come from those exchanges do not exist.

In one case, an American YouTuber and ASMR[4] artist, Gibi, has been repeatedly targeted by deepfakes and online harassment. This issue became so bad that she had to change her name, move out of her home and be extremely vigilant when revealing any potentially identifying information about herself. She has since discovered several online businesses in which others are profiting from selling her image without her consent in deepfakes and other fake, often pornographic, material. She has even been approached by a company offering to remove the deepfakes of her, at $700 a video.[5] Gibi is far from alone: according to 2019 research from Sensity.ai, of the 85,000 deepfakes circulating online, 96% are pornographic, with over 99% of those pornographic deepfakes being of women.[6]

> These harms have a profound negative impact on one's personhood.

But victims of deepfakes[7] and other forms of online harassment lack meaningful ways to obtain justice. In a 2020 survey that captured the online experiences of 484 women and non-binary individuals in the UK during the pandemic, Glitch UK and the End Violence Against Women Coalition found that 83% of respondents who reported one or several incidents of online abuse felt their complaint(s) had not been properly addressed. This proportion increased to 94% for Black and minority women and non-binary people.[8]

What does it say about our digital humanity[9] when we are unable to protect minorities, women and the most vulnerable, as well as ensure that victims have feasible pathways to justice?

An investigation into digital advertising revealed that companies possess 75,000 individual data points about the average US consumer.[10] This data is collected, sold and monetized by private firms, including technology giants Apple, Amazon, Facebook, Google and Microsoft, to feed marketing efforts and create new products and services. That is to say that personal data[11] keeps the data economy going and growing. In fact, as of 2018, the data economy had a combined market valuation of nearly $4 trillion.[12]

The quantity and availability of one's personal data that is available across the data ecosystem is becoming increasingly susceptible to a variety of misuses, which artificial intelligence-driven (AI) systems further amplify. These include intentional misuse of information, such as defamation,[13] misrepresentation or infliction of emotional distress,[14] as well as data-driven discrimination (the limitations of such actions are discussed in Part 3 below). When decisions are made by algorithms, the likelihood and impact of biased or erroneous results can be significant.

Take the issue of facial recognition as an example. According to recent findings conducted by researchers Joy Buolamwini and Timnit Gebru, leading facial recognition software packages performed much worse at identifying women and people of colour than when classifying male, white faces.[15] The 2020 case of Robert Julian-Borchak Williams is the first documented example in the US of someone being wrongfully arrested based on a false hit produced by facial recognition technology. What makes Williams' case particularly important is that then Detroit police chief James Craig acknowledged that the software, if used by itself, would misidentify cases 96% of the time.[16]

These harms have a profound negative impact on one's personhood.

"Personhood" is used in the traditional world to mean recognition of an individual or entity as having status as a person.[17] In the context of this paper, this definition is extended to cover a person's status in the digital world, as well. It is critical to clarify here, however, that in the context

> What happens in the digital world extends to the physical world, and what often seem like violations of privacy are in fact violations of self-determination and personhood.

> Protecting digital rights must thus encompass all forms of self-determination and personhood.

of understanding digital harm, it is not useful to draw false binaries between a "physical" self and a "digital" self. A technology-enabled harm may occur online, but the resulting harms persist regardless of digital and physical borders, and thus do not change the status one still has as a person.[18] Following on from the previous example regarding facial recognition software, the misuse of this technology by law enforcement across the United States has led to several cases of Black men being falsely accused of crimes that they did not commit. These false accusations aren't just an invasion of the falsely accused individual's privacy, they are also an invasion of their self-determination and personhood, affecting their livelihood and reputation.[19] What happens in the digital world extends to the physical world, and what often seem like violations of privacy are in fact violations of self-determination and personhood.

Securing the protection of self-determination and personhood involves acknowledging that digital rights refer to the application of *all* human rights and should not be limited to a number of civil and political rights.[20] Discussions about digital rights are often limited to the rights to privacy, data protection and freedom of expression. The protection of these rights is indeed particularly relevant because, through the use of digital tools, so many aspects of our lives are being tracked, our personal information is being collected, used and disclosed on a massive scale, digital identities are being created, the use of surveillance technologies is spreading and online violence and harassment are increasing.

But the use of digital technology has an increasing impact on all human rights. Tools used to monitor protests and workers' unions affect people's freedom of assembly; algorithms deployed with the aim of predicting and influencing people's behaviour affect their freedom of thought; other automated systems reinforce existing racial or social biases, thus institutionalizing discrimination and other harms. Protecting digital rights must thus encompass all forms of self-determination and personhood.

Part 2

What is digital injustice?

The primary focus in this section is on justice in relation to redress and correction.

> Corrective injustices occur, for example, when the extent of harms is underappreciated or they go fully unnoticed, when there is no accountability or when there are no effective pathways for redressing harms to individuals or groups.

In general, digital justice, in the corrective sense, concerns the rectification of past wrongs – that is, harms that have already been done to an individual or a group. Often, such correction is carried out by imposing a correlative liability on the responsible wrongdoer.

Corrective injustices occur, for example, when the extent of harms is underappreciated or they go fully unnoticed, when there is no accountability or when there are no effective pathways for redressing harms to individuals or groups.

Corrective justice can be used for various purposes, including exploitation and infliction of reputational harm. A recent development in the area of deepfakes is the emergence of a new and far more advanced "nudifying" website that uses deep-learning algorithms to strip women's clothes off in photos without their consent. Anonymous users can upload a photo of a fully clothed woman of their choice, and in seconds, the site undresses them for free. Due to that one feature, it has amassed more than 38 million hits since the start of 2021 and has become one of the most popular deepfake tools ever created.[21] Photos and links from this new website aren't confined to the dark web, nor has it been delisted from Google's search engine index. It operates free of any constraints and has since spread across major platforms such as Twitter, Facebook and Reddit, to name but a few.[22] Victims of non-consensual pornography[23] are forced to grapple with the consequences of these images, losing their jobs, relationships and more. It no longer matters if this content is real or not – they create a new reality in which the victim must live.[24]

However, given the critical knowledge gaps and the degree of uncertainty surrounding the lack of accountability involved, it is necessary to acknowledge that even the best attempts to address these harms in advance may fall short. In light of these potential harms, providing individuals with robust pathways to recourse and redress should be vital aspects of how organizations and governments contend with the ethics and governance of data-driven technologies. This attention to retroactive correction does not mean that efforts to anticipate potential harms (e.g. via foresight methods)[25] and devise proactive guards against them (e.g. by enhancing privacy or fairness by design) lack importance.

In addition, providing individuals with the means of recourse and redress also supports other important social values. For example, researchers have highlighted the importance of due process[26] and corrective justice[27] in relation to the values of rationality and accountability. That is, these mechanisms enable individuals to plan rationally for their lives, knowing they will be protected against unanticipated external interferences and that potential wrongdoers will be held accountable.

What is needed in conjunction with proactive efforts, therefore, is the development of a robust framework for recourse and corrective justice that can support retroactive identification of harms, allocate responsibility and offer equitable pathways of redress. Doing so, however, requires addressing a number of critical challenges.

One challenge pertains to characterizing potentially new types of harm that might be caused by data-driven technologies. Take, for example, the ways in which AI systems undermine LGBTQI+[28] identity. The use of automated gender recognition (AGR) technologies has been shown to remove an individual's opportunity to self-identify, and instead infers their gender from data that is collected about them. This technology uses data such as the person's legal name, whether or not they wear make-up or the shape of a person's jawline or cheekbones to reduce an individual's gender identity to a simplistic binary.[29] This represents a form of erasure for people who are trans or non-binary that, in effect, has dire real-world consequences. When an individual and their respective community is not represented, they lose the ability to advocate effectively for their fundamental human rights and freedoms. Misgendering is particularly harmful due to the sensitive nature of gender dysphoria, and AGR systems further exacerbate the emotional distress associated with an individual's experience with their gendered body or social experiences.[30] Elsewhere, data-driven technologies can have an indirect, but still critical, impact. The influence of recommendation algorithms on shaping

connectivity in social media is an example of this indirect role. Here, the patterns of social connection enabled by algorithms can restrict access to opportunities (e.g. not hearing about a job opening via one's social network).[31]

Another challenge stems from the fact that most moral and legal frameworks depend on clearly identifying the responsible wrongdoer, who is charged with restoring the victim's dignity and reputation in some way or compensating the victim for the harm. Yet, in many cases involving data-driven and predictive analytics, such allocations of responsibility are complicated by online anonymity, the large number of actors involved in the process and the so-called "problem of many hands".[32] Furthermore, even when there is a clear entity to which one could attach responsibility and blame, establishing the fact and degree of harm – even retroactively – can be challenging. The difficulty here is most salient in cases of algorithmic opacity – that is, cases where the grounds on which a predictive algorithm makes or recommends a decision remain inscrutable to humans, including the users and developers of the algorithm.[33] This lack of transparency poses a challenge to the identification of potential harms. The situation is exacerbated by the fact that many of the technical approaches used to make such algorithms interpretable remain fragile[34] and unreliable, particularly in ways that can unhelpfully conceal discriminatory decisions,[35] or else insufficiently grounded in particular aims (e.g. recourse) for which interpretability is sought.

Finally, even if we conceivably possessed the technical means to (retroactively) identify harms, the existing practical roadblocks to redress need to be considered. Such roadblocks might include imposing an undue burden on victims in providing evidence of harm, which is particularly troublesome when the institutions accused of the harm are gatekeepers to this evidence, thus amplifying existing inequities and power asymmetries.[36]

Part 3

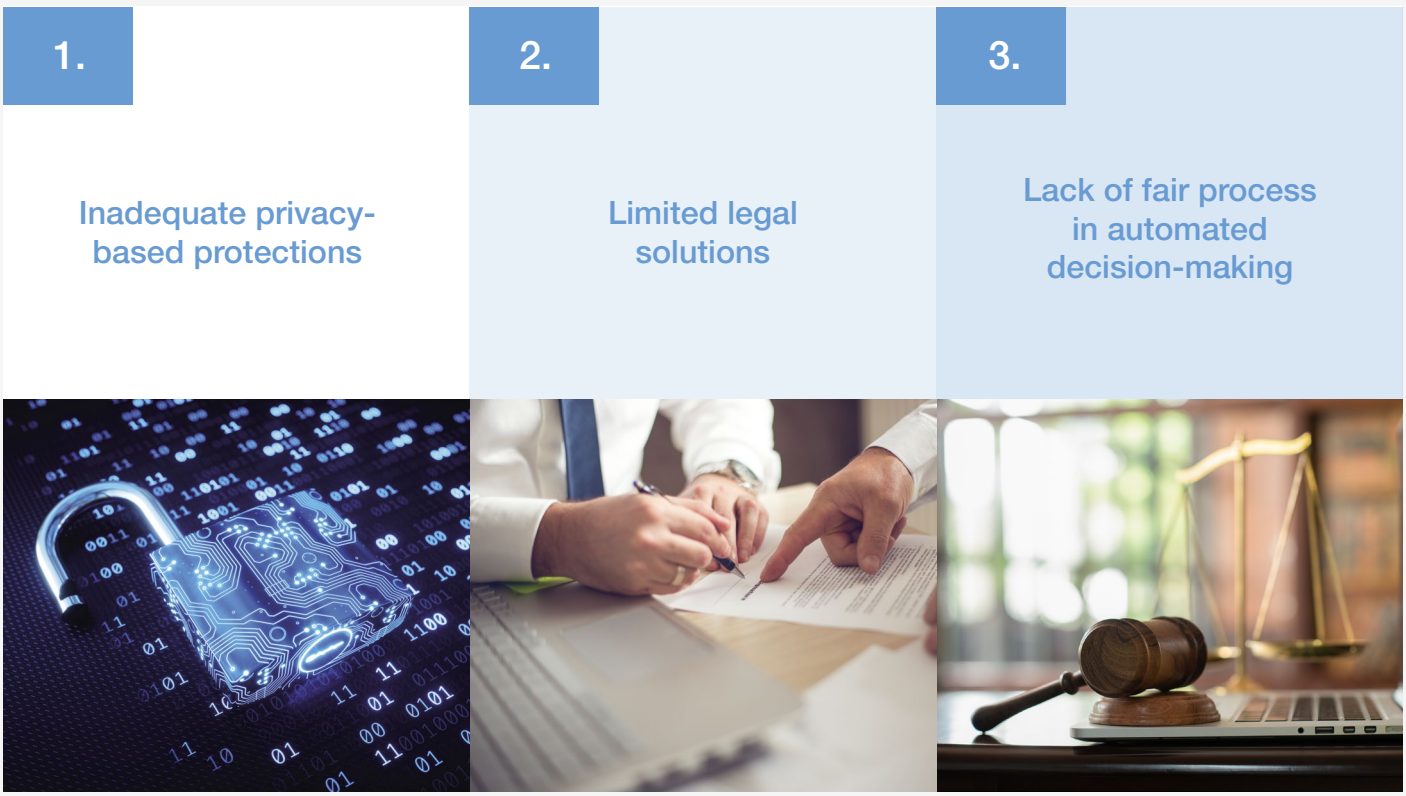
Current legal and judicial systems are fragmented

Emerging technologies pose prominent legal and judicial challenges.

The value of digital tools for human rights and development is enormous. They can help people communicate securely around the globe thanks to end-to-end encryption and support people by increasing their access to information and knowledge. But promoting and protecting digital rights – which must be at the centre of states' strategies for a sustainable digital transformation – requires not only encouraging innovation and technological development but also preventing harms arising from new ways of communicating and analysing data.

Yet so far the protection of human rights – both offline and online – is critically lacking in terms of application and implementation by governments worldwide.[37] Emerging technologies pose prominent legal and judicial challenges, in particular: how to incorporate responsible use of new data-driven innovations into protections that encompass rights beyond just data privacy;[38] how to replace outdated legal codes with frameworks that are fit for purpose in the current digital era; and how to address the lack of fair process in automated decision-making within the context of justice systems.

FIGURE 4 **The three main obstacles to digital justice**



> For these harms to be addressed, states must not only recognize and protect the application of human rights through binding frameworks but also put in place independent oversight mechanisms to monitor the application of these rights.

Arguing that online harms are a matter of privacy has indeed led to notable landmark legal efforts. In the US, the Violence Against Women Reauthorization Act in 2021 (which passed the House of Representatives and has been received in the Senate) includes the SHIELD Act[39] (Stopping Harmful Image Exploitation and Limiting Distribution Act) as an amendment, which proposes to criminalize distributing and intentionally threatening to distribute non-consensual intimate visual depictions of an individual, punishable by two years' imprisonment and a fine. In the United Kingdom, a recent Online Safety Bill[40] was published proposing a new statutory duty of care for social media companies towards its users, including a duty to undertake an "illegal content risk assessment" encompassing terrorist propaganda, child sexual exploitation and abuse content, and other forms of illegal content.[41]

Despite this progress, legal and judicial systems should not treat every technology-enabled harm and abuse as a matter exclusively of privacy, but also as a matter of self-determination and personhood. The United Nations Human Rights Office of the High Commissioner's International Covenant on Civil and Political Rights states that all peoples have the right of self-determination. By virtue of that right they freely determine their political status and freely pursue their economic, social and cultural development.[42] Rights to self-determination personify the right to live and prosper in whatever way one deems fit, *beyond* the notions of who or what can process their data and for what purpose.

It is a commendable start that some countries are introducing domestic legislation to tackle digital harms; however, its effectiveness is limited by jurisdictional issues as well as the scale of sophistication that many technology-enabled harms, overall, are taking.[43] Both challenges will need cross-border collaboration and universal data protection standards that prioritize people's fundamental human rights.

For these harms to be addressed, states must not only recognize and protect the application of human rights through binding frameworks but also put in place independent oversight mechanisms to monitor the application of these rights.

A Inadequate privacy-based protections

> The global cost of data breaches in 2021 alone is expected to reach $6 trillion.

Data has undoubtedly helped improve products, services, policy responses, studies and investigative journalism. But data, and personal data in particular, can also be used for a variety of harmful purposes that affect both individuals and wider society. The global cost of data breaches in 2021 alone is expected to reach $6 trillion. Moreover, rich datasets and novel models are used to track and repress minorities in many countries across the world, unconstrained by any rules or oversight. It is no wonder, therefore, that people increasingly report high levels of concern about the lack of data privacy and security.

These concerns have led to the adoption of various laws and regulations to limit the misuse of data as well as to incentivize better data management practices. The European Union famously opted for the General Data Protection Regulation (the GDPR),[44] one of the most wide-ranging pieces of legislation passed by the political bloc and market of hundreds of millions. It levies harsh fines against those who violate its privacy and security standards; notably, under certain conditions, it also applies to companies that are not in Europe. In China, the Personal Information Protection Law (PIPL) was adopted by the Standing Committee of the National People's Congress on 20 August 2021, and it will come into effect on 1 November 2021.[45] In the US, while there are no privacy laws at the federal level, some vertically focused federal privacy regulations and consumer-oriented state laws such as the California Consumer Privacy Act have sought to limit abuses.[46]

However, even the most effective of these have failed to keep pace with the speed at which technology evolves, leading to perpetual mismatches and deficiencies. Legislation has done little to address the lack of transparency in the inner workings of algorithmic processing of data, further obstructing the path to accountability and scrutiny. This is illustrated by the numerous reports of algorithmic discrimination; for example, in relation to housing markets and job recommendations. Laws also tend to focus exclusively on personal data (often narrowly defined) and may fail to address the way in which data is used to infer and predict characteristics that can lead to harm to individuals and groups even when their own personal data is not collected. In addition, the complexity, rigidity and fragmented nature of privacy legislation has decreased its effectiveness and led to negative externalities, too. Compliance can be costly and difficult, particularly for smaller players.

How can we, therefore, limit the misuse of data as well as incentivize better data management and governance practices that enhance traditional privacy law principles to ensure they are fit for purpose? Solutions lie in proportional regulatory oversight of data-driven decision-making, and the ability to audit the inner workings and impacts of high-risk algorithms and AI on society. Given the highly technical characteristics of some of these systems, new and experimental regulatory paradigms should be explored and trialled. Regulatory markets, middleware solutions,

interoperability requirements, auditing mechanisms, transparency obligations and accelerated group litigation are directions to explore.

Lastly, an enabling environment should be defined. Legislative alignment and the facilitation of cross-border data flows while ensuring rights protection and security should help maintain a competitive ecosystem. Decreasing barriers to entry and incentivizing harmonization through a modern, flexible and enhanced data privacy standard may facilitate international trade between responsible actors while safeguarding the primacy of fundamental rights.

B Limited legal solutions that are no longer fit for purpose

Initial government responses to technology-enabled harms often focused on protecting the democratic process from political manipulation and disinformation, and more recent responses have also focused on countering COVID-19 disinformation. However, governments need to protect not only themselves but also their citizens. Individuals who have been harmed have become increasingly adept at using existing laws – ranging from defamation and invasion of privacy to copyright infringement and unfair competition – to sue their attackers directly. Most such causes of action predate the internet, though, and while they can be pressed into service, they are not necessarily well-adapted to this function, particularly when confronted with defendants who might be anonymous or located outside the legal systems' reach.

In terms of criminal law, offences such as harassment, designed for an offline world, can be pressed into service against online activity, and many countries[47] have adapted this offence to apply more easily online. Early specific legislation[48] provides for broad offences, but these definitions are often clunky, overbroad and ill-suited to

> Early specific legislation provides for broad offences, but these definitions are often clunky, overbroad and ill-suited to more current forms of data-driven harms.

more current forms of data-driven harms. Recent legislation is more appropriately targeted at specific harms, such as the non-consensual recording, distribution or publication of intimate images.[49]

To address this, governments should first consider modifying existing laws or crafting new ones to permit appropriate private and criminal redress for data-driven harms, and to discourage harmful behaviour in the first place. One area in which governments are moving forward is in respect of privacy and data protection;[50] but governments must not treat every technology-enabled harm as a matter solely of privacy. Hence, much more needs to be done, not only in respect of privacy (see Part 3A above), but also more generally regarding the technology-enabled harms discussed previously (see Part 2).

At the dawn of the internet era, governments enacted legislation to ensure that legal considerations did not inhibit the growth of the internet, including safe harbours[51] and immunities,[52] which meant platforms were not liable for damages for their users' content and actions. This policy has been extraordinarily successful for the platforms themselves, but it has also enabled the rise of harmful digital content. The second thing governments must do is consider whether the balance that was appropriate several internet generations ago is still appropriate now. Potential changes range from refinement (so that certain kinds of harmful content no longer fall within the safe harbour or immunity),[53] through various significant reforms (so as to increase the obligations on very large online platforms,[54] or enlarge the range of content that can be subject to takedown,[55] or impose short time limits for doing so)[56] to outright abolition.[57]

> Moreover, pathways to restoration for individuals and communities affected by technology-enabled harms must extend beyond the scope of what is currently considered possible under the law.

In fact, governments (such as France,[58] Germany, [59] India[60] and Russia[61]) have already legislated in this space, and others (such as Canada,[62] Ireland[63] and the UK[64] – as well as the EU[65]) are about to do so. Growing impatient with platform self-regulation and going beyond the kind of co-regulation exemplified by the EU's Code of Practice on Disinformation, governments are considering the imposition of wide-ranging regulation, including broad definitions of harmful digital content, additional duties of care for platforms, rapid compelled takedowns, mandatory content moderation and website blocking, criminal sanctions and powerful regulators. One influential model for this is provided by the Australian eSafety Commissioner.[66]

The third thing governments must do is ensure that legislation in this space broadly converges on international standards of effective methodologies – especially in relation to cross-border actions. It bears repeating that governments enacting such legislation should ensure that definitions of harmful digital content take into consideration, in particular, the technology-enabled harms discussed above (see Part 2).

However, changing the legal regime surrounding data-driven harms risks replacing one set of blunt instruments with another, or replacing an overly tolerant legal regime with one that encourages or even mandates censorship. Care must be taken here, and good intentions must not be pressed too far. Overbroad restrictions on online communication have been struck down in the US,[67] India,[68] France[69] and Germany[70] – as well as in the European Court of Human Rights.[71] Opponents to rolling back platform immunities argue that even well-intentioned reforms would harm marginalized individuals and communities, sometimes the very persons the reform was intended to protect.[72] It is difficult to devise reasonable periods and fair procedures for compelled takedowns (see Part 3C below), scalable content moderation mechanisms, workable website blocking procedures, narrowly drawn and practicable obligations to provide user data to law enforcement agencies, realistic transparency obligations, appropriate criminal sanctions, powers of regulators and carefully drafted definitions of harmful digital content (though, in this last respect, the principles to do so are discussed above [Part 2]), particularly if those rules are intended to cover a wide range of entities with widely differing purposes, interests and capabilities. Conversely, it is very easy for such matters to expand well beyond their legitimate scope. Consequently, the fourth thing governments must do is ensure that legislation in this space does not go too far and prioritizes elements of fairness in order to ensure justice, legitimacy and inclusion. Moreover, pathways to restoration for individuals and communities affected by technology-enabled harms must extend beyond the scope of what is currently considered possible under the law.

C Lack of fair process in automated decision-making

“ We are now undergoing a technological transition in how societal decisions are made and, in turn, realigning shared normative expectations for what constitutes fairness in human-computer interactions.

As decisions are increasingly influenced or even exclusively made by algorithms, the preservation of fair process has become an urgent matter. This is especially urgent within the context of ensuring sufficient redress mechanisms for victims of data-driven harms, and preventing bias, discrimination or inequitable outcomes. For these reasons, preservation of fairness is key to ensuring justice, legitimacy and inclusion, and thus requires features of consistency of application, accountability and transparency.

We are now undergoing a technological transition in how societal decisions are made and, in turn, realigning shared normative expectations for what constitutes fairness in human-computer interactions and digitally networked transactions. What is at stake in this transformation of social norms is the affirmation of fundamental human rights and civil liberties governing how rules are conveyed, what data is evaluated to make judgements, when exceptions and limitations are applied, whether there is an opportunity for appeal and a host of other revived fair-process issues in the context of automated decision-making. These hard-fought protections that have evolved over a long history of emancipation and freedom need to be defended again.

Civil and criminal procedure may vary by jurisdiction, and administrative proceedings are conducted differently across legal systems, but the expectation of fairness in process has always been innate to human behaviour. The recent adoption of video-assisted reviews of referees’ calls (football/soccer has video assistant referee technology [VAR], American football uses instant replay, etc.) was designed explicitly to eliminate human error and address fair process complaints about on-field referees/umpires with the aim of improving consistency, providing accountability (referees/umpires may see their decisions overturned) and increasing transparency (coaches and fans can see the same video as the reviewer). As the significance of automated decision-making in legal and judicial systems increases, agreement concerning fair process becomes more essential for the legitimacy of the social order.

But the immense quantity of digital communications and data can make fair process difficult to achieve. Along with these huge networks of increased access to services, lower prices and faster service has come a volume of questions, disputes and complaints that would be unmanageable using existing structures of decision-making in the analogue world. The implementation of automated processes to manage this volume using increasingly sophisticated algorithms is inevitable – and does in fact demonstrate the potential for technology to lower barriers to entry and result in greater inclusion.

However, automated decision-making often fails to satisfy standards of fair process. There are many situations in which the same rules would apply in the same way given very similar circumstances so that an automated decision could be consistently applied. But there is also a large swathe of scenarios that require human judgement and a certain degree of empathy to be adjudicated fairly.

In addition, fair process includes the expectation that decisions should be transparent and subject to appeal – but an algorithm's answer can often be unassailable, either because there is no mechanism to challenge it or no explanation has been given for how it reached its answer.

Research into algorithmic risk assessment provides one telling example of how the harmful, unintended consequences of these technologies on individuals, particularly members of marginalized and Indigenous communities, proliferate due to the lack of regulations that ensure lawfulness, fairness and transparency.

Back in 2016, journalists working with non-profit newsroom ProPublica investigated concerns being raised by various communities regarding the use of the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) algorithm, which produced scores allegedly predicting the likelihood that people charged with crimes would commit future crimes. These scores were presented to judges as relevant information used to determine sentencing and billed as a valuable tool to prevent repeat offences.

However, the ProPublica investigation determined that COMPAS "proved remarkably unreliable in forecasting violent crime". The journalists reviewed risk scores assigned to 7,000 people arrested in Broward County, Florida, in 2013 and 2014 and found that only 61% of those deemed "high risk" to commit future crimes actually did so. Worse, they found that the system was highly biased: it was much more likely to flag Black defendants as liable to commit further crimes in the future, wrongly labelling them at twice the rate as white defendants. White people were also wrongly labelled as low-risk more often than Black defendants.[73]

As a result, sentencing decisions based on COMPAS may lead to disproportionately harsh sentences or denial of parole for Black defendants. In one particularly egregious case, Paul Zilly, a Black Wisconsin man, agreed to a plea deal offered by the prosecution, but the judge rejected the plea deal and imposed a new sentence that doubled Zilly's time in prison based in part on a COMPAS determination that Zilly was a high-risk offender.[74] Rather than accepting a negotiated plea deal, the judge relied on a prediction that was not subject to any professional oversight, validation or accountability.[75]

If computers could accurately predict which defendants were more likely to commit new crimes, the criminal justice and social services systems could work towards prevention and/or more tailored forms of redress. The challenge, however, is to ensure that the computer gets it right – which has yet to be achieved.

New approaches are needed that merge algorithmic efficiency with the fundamental elements of fair process. There are persistent aspects of basic criminal and civil procedures for administering regulations across legal processed – such as the issuance of notice and being afforded the opportunity to correct inaccuracies and challenge matters of dispute, with an impartial judge determining the outcome. There should be technical structures built into the algorithm to support the capacity to contest a decision, with a predictable and explainable process to administer it. A notion of proportionality can help guide the application of fair process, but its fundamental principles should be universally upheld. As these algorithms learn over time, the assessment for risk of bias needs to be revisited to ensure that fair process is maintained.

Part 4

Recommended pathways to digital justice

Pathways to digital justice must be victim-informed and protect all forms of self-determination and personhood.

No matter how effective the initial design of a digital system, it is bound to need pathways to justice – this means unobstructed access to human rights. Pathways to digital justice should provide clear and feasible steps for any individual or group to obtain justice for whatever form of harm has been inflicted. Most importantly, in order for a pathway to effectively lead to justice, the corrective steps and mechanisms developed must be victim-informed and encompass the protection of all forms of self-determination and personhood. These factors were considered in deciding upon the two subsequent recommendations articulated here: (1) to modernize the capacity of the judicial system in order to adjudicate more claims; and (2) to equip survivors with timely, feasible steps to navigate the justice process in the event of some form of harm.

FIGURE 5 **Two multistakeholder recommendations to create clear pathways to digital justice**

Recommendations	Examples
1. Increase judicial system's capacity in order to adjudicate more claims	Look to existing experiments in private- and public-sector spaces such as eBay's online dispute resolution system, which has been adopted by governmental and private adjudication systems
2. Develop timely and feasible steps for victims to navigate the justice process in the event of some form of harm	Create a victim resource guide with at minimum the 10 core victim-centred components as outlined in the World Economic Forum's digital justice paper

A Increase systems' capacity to adjudicate more claims

Platforms and governments can and should do more to address and redress the full range of data-driven harms. In particular, governments should remove unnecessary obstacles to existing causes of action and craft better-targeted civil and criminal offences; they should reconsider whether legal and policy choices that were relevant several internet generations ago are still appropriate now; they should ensure that much-needed legislation in this space broadly converges with international standards of digital rights and related areas; and they should ensure that legislation in this space prioritizes fair process in automated decision-making – especially within the context of the justice system.

As digital transformation continues to have an impact on a broad range of rights-affecting systems, there has been a growing recognition of its potential to create novel rights and harms. What has received less attention is that the development of digital systems also creates the opportunity to build novel mechanisms to realize those rights and resolve the disputes that will inevitably arise from their exercise. In some circumstances, there may be a need to build systems that provide people with what scholar Danielle Citron has called "technological due process",[76] ensuring that the baseline procedural rights intrinsic to government processes are not overlooked or even actively violated as those processes undergo digital transformations. This highlights an even larger issue: the need for accessible, enforceable procedural protections across a wide range of human rights in digital systems, corporate as well as governmental.

It is worth recognizing that the majority of digital justice dialogues, including the one in this paper, centre on reactive approaches to realized harms. There is an opportunity, however, to centre digital transformation initiatives and system design around strengthening and protecting specific, articulable rights. Systems prioritizing this frame would design for many of the same agency and redress rights, but measure success on impact, as opposed to volume.

One of the largest challenges in building pathways to digital justice is that they are often based on analogue pathways to justice. But the globally connected nature of many technologies has implications for the design of a related justice system. If a user in one country, for example, is wronged by a person in another country, on a platform based in a third country that promises to prevent that harm in another country, where does a person even start? And why is figuring out the mechanism of justice the individual's responsibility? Even if these questions were simply answerable by "bring a case to your local court" (which it is not in most places around the world), the fact remains that most formal systems of justice are practically inaccessible to most of the people they serve. Even if we solve the backlog of cases and access to justice issues, the truth is that traditional justice systems have a limited ability to meaningfully adjudicate, let alone fix, the kinds of problems that digital systems cause.

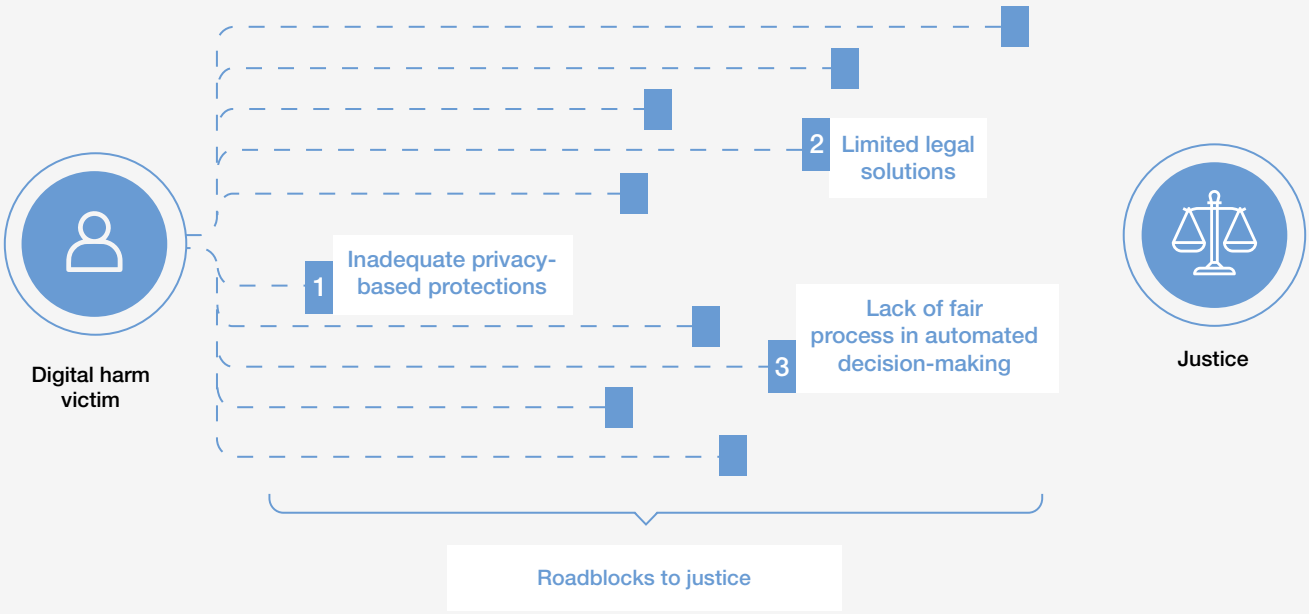
Take, for example, three different approaches to justice and redress:

- **Corrective:** How does a person whose rights are harmed compel the actors involved, to the extent possible, to undo the offending action? (See Part 2 for a detailed explanation)
- **Restorative:** How does a rights holder recover, to the extent possible, from the impact of the harm?
- **Punitive:** How does a rights holder hold the person responsible accountable?

Governments and platforms have an opportunity to design digital justice systems capable of addressing these approaches, but the issues quickly become complicated. How, for example, do you compensate someone for the impact of a deepfake? Who should be held accountable for software or algorithmic bias errors in healthcare benefits or bail recommendation systems, and what would that accountability entail? How does a platform company fairly balance user expectations when the users come from conflicting justice traditions?

In reality, these questions are being answered by platforms all the time; however, the systems are not usually in place to make those answers consistent or binding. A substantial amount of digital justice work takes place on the continuum between platform customer service and formal legal or regulatory action. And each has an important, ideally complementary, role to play. When it comes to designing pathways to justice, often the hardest part is not achieving justice, but how accessible, participatory and interoperable the pathways are.

FIGURE 6 **Victims of digital harm have no pathways to justice**



Rather than suggesting simple or singular answers, this paper reviews a few examples that illustrate key issues for designing pathways to digital justice:

- **eBay's online dispute resolution system**
  When eBay launched its groundbreaking online marketplace, most people didn't trust ecommerce. Recognizing the need to establish trust, eBay's team developed the world's largest online dispute resolution (ODR) system. At its peak, the ODR system resolved more than 60 million claims a year, supported by a staff of 12. ODR as a methodology has been adopted by governmental and private adjudication systems. eBay's example illustrates the efficiencies and market-building that proactively designing pathways can offer to digital platforms and marketplaces.
- **UK A-levels replacement**
  The Government of the United Kingdom, in an effort to standardize student qualification scores without the ability to administer their traditional A-level exams during the COVID-19 pandemic, developed an algorithm to predict exam results. The algorithm reflected predictable biases,

> Courts are where the balance between state power and individual rights is at its most delicate; the way in which legal systems not only build pathways to justice for the victims of deepfakes but also ensure that deepfakes do not become a weapon of injustice is a critical frontier for the next generation of digital transformation.

favouring the already privileged, and became the subject of popular outcry and legal threats by digital rights advocates. The government, in response to the pushback, abandoned the algorithmic inputs and went, instead, with teacher evaluations. The end result was relatively popular, but the pathway to justice in this case would have been best established if there had been more advocacy directed to the UK government against the use of this kind of data in the algorithm, as well as against the underlying inequities that provide the data that was fed into the algorithm in the first place.

– **App-based ride services**
There are few industries that have challenged as many norms and legal frameworks as app-based ride services such as Uber and Lyft – and, in some ways, their litigation history is a roadmap to unmet data governance demands in several types of relationships. For example, there have been several driver-led class-action lawsuits, challenging aspects of the digital transformation of the employer and employee/contractor relationship. Similarly, while most platforms have sophisticated digital customer service processes, there are a number of customer complaints that they can't answer. For example, some app-based ride services acknowledge thousands of complaints of sexual harassment and assault with no meaningful organizational reaction. As a result, Uber and Lyft are both defendants in a large number of legal cases alleging that drivers sexually assaulted riders. In both examples, the courts are being used to challenge the absence of pathways to justice for predictable harms, happening at scale in digital transportation companies.

– **Deepfake vs. justice**
In the US, there's a growing recognition of the accessibility and harms of a wide variety of impersonation technologies, but the development of policies to mitigate those harms is just beginning. One example is in digital evidence law – essentially, when and how data and digital representations can be used to influence the outcome of court proceedings. Court systems use evidence law as a way to ensure the integrity and provenance of information that influences rights-affecting decisions, but most court systems are still developing both the legal and technical capacity to address deepfakes and other digitally created or altered content. Courts are where the balance between state power and individual rights is at its most delicate; the way in which legal systems not only build pathways to justice for the victims of deepfakes but also ensure that deepfakes do not become a weapon of injustice is a critical frontier for the next generation of digital transformation.

These case studies illustrate a range of digitally intermediated pathways to justice and, in some cases, their absence or conflict with institutional mechanisms of justice. Ultimately, digital governance and justice are emerging concepts – not only at the substantive and political level but also at the mechanistic and procedural level. The nascency of political consensus on digital justice also suggests the value of building public experimentation, validation and subsidization infrastructure that tests the participation structures driving the best outcomes.

While there are likely to be many models that succeed, based on contextual factors, there are also high-level methods to evaluate the maturity of a pathway to justice. Here, maturity implies the availability of a certain level, stage and sophistication of participation tools to enable rights holders to realize those rights.

- **Articulation and transparency**
  A digital system recognizes its potential to affect rights in a particular way and creates vehicles to provide transparency in its decisions.

- **Education and engagement**
  Those responsible for a digital system proactively use their points of contact with users to create awareness of the rights impact of their work and to create direct reference resources that enable users to realize those rights.

- **Reporting**
  A digital system has an embedded process that help users identify and raise the profile of incidences and/or types of harm.

- **Accountability and redress**
  A digital system tracks the provenance of a line of service and/or rights-affecting behaviour, enabling rights holders to directly and specifically hold abusers accountable, as well as to seek redress for any harms experienced.

- **Escalation to independent oversight**
  A digital system is directly and transparently subjected to relevant jurisdictions and authorities, both as a direct escalation of complaints insufficiently managed by self-regulated processes and as a mechanism to align incentives with rights holders.

These are, of course, the early stages of building pathways to digital justice – but each stage represents a scale change in the maturity of participation models towards establishing legitimate, rights-protecting digital justice.

B Create a victim resource guide

At a practical level, every government, globally, should maintain a digital justice victim resource guide that is easily accessible by everyone. Individuals and communities need clear steps to justice in the event that they experience harm. Victims should not have to navigate countless channels to address a single digital crime, particularly when time is a major factor.

Below are 10 core components that should be included in a digital justice victim resource guide to ensure that it will be helpful to victims:

1. A summary of existing national laws that apply to data-driven harms

2. A clear roadmap detailing what can and cannot be realistically done under the current law to: 1) stop digital crime; 2) hold the bad actor accountable; 3) remove harmful content online; 4) prevent future abuse; and 5) provide victims with effective compensation

3. Sample language that victims can use to explain their situation to the institution from which they are seeking help

4. Common questions that victims may be asked by a caseworker or investigator, to help them prepare for the traumatizing experience of communicating an abusive event. This should include specific guidelines about what information victims need to provide

5. Name and contact information for the departments that provide services to victims of digital harm. The recommendation is to enlist a “digital victims investigator”, who is mandated to work directly with digital abuse victims, has experience working with both victims and technology, is able to work across jurisdictions and is trauma-informed

6. Guidelines on how victims can track and monitor their claims. It is recommended to create a digital justice investigation portal that victims and their digital victims investigator can access to receive up-to-date information on their matter and communicate with staff assigned to their case

7. Expectations and next steps for victims after they have gone through steps 1–6

8. List of help hotlines for victims of digital harm

9. List of support groups for victims of digital harm

10. Clear, achievable preventative actions that any victim can take to better protect themselves from future potential harms

For the developing world, more foundational elements will need to be established to help victims with lower digital literacy or governments that are in the earlier phases of the digital transformation journey. This point was highlighted in the World Bank's 2020 report about investing in the digital literacy of people to enable active and informed participation, stating that, even in a policy environment with strong protections for individuals' data, people must also have the requisite skills and awareness to engage in the data ecosystem actively and responsibly.[77]

Creating a digital justice victim resource guide is not the be-all and end-all solution, but it is a productive step in the right direction. This guide should apply not only to citizens of a respective country but also to tourists or visitors who may experience digital harm in a different country. This could promote cross-border collaboration and push for international cooperation in this space to break past territorialism.

Furthermore, as important as it is for every government to create a victim resource guide, it would be similarly beneficial to provide clear steps on how to seek redress, articulated at the technical and industry-wide level. At the end of this white paper is a sample victim resource guide geared towards deepfakes, and a sample tip sheet for governments and law enforcement agencies, which offers high-level insights that can inform a victim-oriented digital justice process, globally.

FIGURE 7 **What new pathways for digital justice should include**

Pathways to digital justice must be:

1. Accessible
2. Interoperable
3. Participatory
4. Trauma-informed
5. Victim-centred

Conclusion

No legislation, regulation or policy can truly be effective at protecting the fundamental human rights of individuals, particularly women, LGBTQI+, BIPOC[78] and other historically marginalized communities unless it creates accessible, participatory and interoperable pathways to digital justice. This paper has recommended a variety of mechanisms for corrective justice that can help halt the increase of data-driven harms, as well as proactively tackle this issue at a foundational level. There are several civil-society organizations, researchers and activists that have all, in different ways, also identified the need for these changes to happen. The hope is that this paper contributes to existing scholarship and notable ongoing progress in this area, as well as underscoring the reality that, ultimately, governments have the duty and responsibility to lead this urgent task to enforce human rights and privacy laws that are fit for purpose in the current digital era, that close the accountability gap, and that include fluid cross-border redress mechanisms, strong fair process and trauma-informed judicial processes.

Doing so should muster the cohesive, multistakeholder action required to combat a potential next generation of digital injustices.

WORLD ECONOMIC FORUM

DIGITAL JUSTICE
Tip Sheet

FOR LAW ENFORCEMENT AND TECH PLATFORMS

> THE NUDE IMAGES WERE OFTEN ACCOMPANIED BY IDENTIFYING DETAILS INCLUDING VICTIMS' ADDRESSES, PLACE OF WORK, LEADING TO THE FEELING OF THREATS TO THE VICTIMS' PHYSICAL SAFETY.

PURPOSE

This document offers insights to aid policy-makers in creating trauma-informed policies that law enforcement and platforms can follow to approach digital justice in a victim-oriented way.

TECH-ENABLED TRAUMA

Suggestions to avoid unintentional re-traumatization:

Instead of requiring a victim to re-tell (and re-live) their story, try creating an initial comprehensive report.

Instead of asking a victim "Why did you send the intimate photo?", try focusing on the actions of the online attacker(s).

Instead of suggesting that they make their accounts private/get off social media, try asking victims what would make them feel safer online.

> WREN (AGE 14) HAD TO BE REMOVED FROM SCHOOL DUE TO BULLYING AND SHAMING IN HER COMMUNITY. SHE NEVER PURSUED LEGAL ACTION.

INSTITUTIONAL TRUST

Cooperation and trust between law enforcement and tech platforms is ***critical***. Without it, investigations, regardless of the laws/policies implemented, will stall and victims will suffer.

EXPLOITATION

Image-based sexual abuse (sextortion, non-consensual pornography and deepfake pornography) is a devastating form of online expoitation.

ACTING EARLY

Online harms proliferate rapidly, which underscores the need for early intervention by law enforcement and relevant tech platforms.

ABUSE TYPES

Other forms of online exploitation include:

Non-consensual tracking, online harassment, cyber bullying, doxxing, impersonation, ransomware/digital extortion schemes, social discrimination, algorithmic bias/paternalism and others.

DIGITAL VICTIMS INVESTIGATOR

A suggested position that works directly with online abuse victims, understands tech, works across jurisdictions and is trauma-informed.

ENDTAB.ORG

DEEPFAKES

a victim resource guide

GUIDE OVERVIEW

This guide includes resources, tools and strategies for anyone who has been targeted by a **deepfake**, or similar technology, which has made it falsely appear as if they were in a nude photo or pornographic video that they were not actually in.

CREDIT: FACEBOOK

YOU ARE NOT ALONE

96% of all deepfake videos are pornographic and almost all of these target women. This is a pervasive new form of online image-based sexual abuse and it is unacceptable. Because this technology is new, some people may not be aware of deepfakes and there may not be laws against deepfake pornography in your jurisdiction.

This means you may need to educate others when seeking help. This guide is designed to help you with that.

1 DEEPFAKE PORNOGRAPHY: WHAT IS IT?

These videos use face-swapping technology to transfer a victim's face from a photo or video onto the body of someone else in a pornographic video, making it falsely appear as if the victim is engaging in sex acts. To safely see how this technology works, click **here** to watch this explainer video created by Facebook.

CREDIT: TIKTOK, DEEPFAKE TOM CRUISE

2 X-RAY & "NUDIFYING" PHOTO APPS

These apps turn photos of famous and everyday women and girls into realistic naked photos ("fake nudes") using deepfake technology to "remove" their clothing. These apps do not work on photos of men.

EXPLAINING THE HARM

WHEN VIEWERS BELIEVE IT'S REAL

When believed, the harm caused by disseminating deepfake pornography or a fake nude is similar to non-consensual pornography. It can adversely affect a person's employment, reputation, relationships and emotional well-being, and upend their life. Never knowing when or if said video or photo will show up can leave victims in a perpetual state of anxiety and fear.

WHEN VIEWERS KNOW IT'S FAKE

Deepfake pornography and fake nudes do not have to be believed to cause harm. Even if labelled as fake, a pornographic deepfake depicting a victim that is shared and watched by others is an act that sexualizes and fetishizes them publicly without their consent. Under any circumstance, virtually forcing a victim to engage in a sex act is harmful.

3 SAVING THE EVIDENCE

Before deleting anything, be sure to preserve any evidence. Some ways to do this: (1) download the video(s); (2) screenshot the photo and webpage, including the url, date and time; and (3) save the webpage as a pdf. Visit withoutmyconsent.org to access its evidence preservation resources for non-consensual pornography (also effective for deepfake pornography).

4 REMOVE FROM GOOGLE SEARCHES

Google enables victims of deepfake pornography (referred to as "fake pornography" by Google) to remove the offending video or photo from appearing in its search results. (It does not mean it is removed from the actual website.) Click here to begin.

INSTAGRAM YAHOO!
FACEBOOK GOOGLE
TWITTER SNAPCHAT
REDDIT TUMBLR
PORNHUB

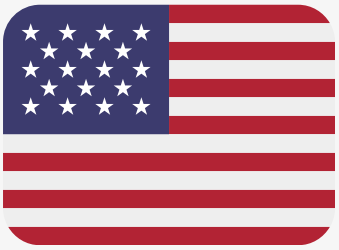
5 TAKEDOWN REQUESTS

Many technology platforms will voluntarily remove posts such as non-consensual pornography because they violate their content policies. Deepfake pornography should be treated similarly. The Cyber Civil Rights Initiative created guides to do this on different platforms. Click any of the sites listed on the left to start the process or, to see the entire guide, visit: www.cybercivilrights.org/online-removal/

6 TAKEDOWN REQUESTS: COPYRIGHT

A deepfake of Kim Kardashian was **removed** from YouTube based on a copyright takedown request from the company that created the original video. It may be possible to apply this approach to deepfake pornography by alerting the publisher of the underlying pornographic video. They may be able to issue a copyright takedown notice. Or you may seek to issue the takedown notice yourself following the same approach used for non-consensual pornography **here**. We recommend speaking with a lawyer if you have questions.

7 DEEPFAKE LAWS



UNITED STATES

States are slowly beginning to pass laws, both civil and criminal, that specifically address deepfake pornography and electoral interference. To date, these states include CA, VA, HI, NY, MD, WY and TX.

AUSTRALIA

Australia prohibits the non-consensual distribution of deepfakes. Under Australia's image-based sexual abuse laws, the non-consensual distribution of deepfakes would be prohibited in Australia at the federal level, and in most states and territories in Australia with image-based sexual abuse laws. For more information, see **this article** by Noelle Martin.

FALSE IMPERSONATION
HARASSMENT
EXTORTION
IDENTITY THEFT
CREDIBLE IMPERSONATION
DOMESTIC VIOLENCE

8 EXISTING CRIMINAL LAWS

When reporting deepfake pornography to law enforcement officers, they may need to use existing criminal laws against the perpetrator. Some suggested laws are listed on the left. *NOTE: Non-consensual pornography (revenge porn) laws will likely not apply because it is not the victim's body (only their face) in the video.*

9 RESTRAINING ORDERS

You may choose to seek a restraining order or order of protection requesting that the restrained party takes down, destroys and does not distribute the deepfake. It is abuse and may qualify as harassment, disturbing your peace, false impersonation, threatening behaviour, causing emotional distress and more. Contact your local domestic violence or sexual assault organization for assistance.

SAMPLE LANGUAGE:

I am the victim of a new form of online abuse called non-consensual "deepfake" pornography. These deepfake videos use face-swapping software or an app to transfer a victim's face onto the body of someone else in a pornographic video, making it appear that the victim was in a video they were not in fact in. [Name of Restrained Party], using only images of my face, created and distributed/is threatening to distribute a pornographic deepfake video of me without consent. It looks like me in the video, but it is not me. [Name of restrained party] has plenty of my images, meaning there is nothing to stop them from continuing to manufacture more fake pornographic videos of me at will.

EXPLAINING DEEPFAKES

Contributors

Lead Authors

Evîn Cheikosman
Policy Analyst, Crypto Impact and Sustainability Accelerator (CISA); Manager, Global Future Council on Data Policy, World Economic Forum

Sina Fazelpour
Assistant Professor of Philosophy and Computer Science, Department of Philosophy and Religion and the Khoury College of Computer Sciences, Northeastern University

Sheila Warren
Deputy Head, Centre for the Fourth Industrial Revolution; Member of the Executive Committee, World Economic Forum

Acknowledgements

Design Strategy

Tricia Wang
Tech Ethnographer, Sudden Compass

Global Future Council

Special thanks to the core drafters from the Global Future Councils on Data Policy; AI for Humanity; and Media, Entertainment and Sport:

Wafa Ben-Hassine
Principal, Responsible Technology, Omidyar Network

Alberto Giovanni Busetto
Group Head of Data and Artificial Intelligence, The Adecco Group

Anna Byhovskaya
Senior Policy Adviser, Trade Union Advisory Committee (TUAC) to the OECD

Juhi Garg
Co-Founder, ED Times

Lesly Goh
Fellow at Cambridge University, Former Chief Technology Officer at World Bank Group

Bret Greenstein
Consulting Principal, Analytics Insights, PwC, Former Senior Vice-President, Global Markets; Head, Artificial Intelligence and Analytics, Cognizant

Ayesha Khanna
Chief Executive Officer and Co-Founder, ADDO AI

Sushant Kumar
Principal, Beneficial Technology Investments, Omidyar Network

Sean McDonald
Co-Founder, Digital Public

Naveen Menon
President, ASEAN, Cisco Systems

Monique Morrow
Senior Distinguished Architect, Emerging Technologies, Syniverse

Marietje Schaake
Director, International Policy, Cyber Policy Center, Stanford University

JoAnn Stonier
Chief Data Officer, Mastercard

Amjed Al Thuhli
Information Technology Consultant, Information Technology Authority (ITA)

Priya Jaisinghani Vora
Chief Executive Officer, Future State

Advisory Committee

Special thanks to members of the Pathways to Digital Justice Advisory Committee:

Sophie Compton
Co-Founder, My Image My Choice

Chris Conley
Former Technology Policy Attorney, The American Civil Liberties Union (ACLU)

Adam R. Dodge
Founder, EndTab

Mary Anne Franks
Professor of Law and Dean's Distinguished Scholar, University of Miami School of Law; President, Cyber Civil Rights Initiative

Karin Gabriel
Head – Future Thinking School and Home Delivery Services, Ars Electronica

Samuel Gregory
Program Director, WITNESS

Eva Kaili
Member of the European Parliament

Noelle Martin
Lawyer, Activist

Eoin O'Dell
Associate Professor of Law, Trinity College Dublin

Giorgio Patrini
Chief Executive Officer and Chief Data Scientist, Sensity.ai

Nina Schick
Adviser and Author, Deepfakes: The Coming Infocalypse

Matthew Turek
Program Manager, DARPA
Information Innovation Office

Special thanks to:

Marcus Burke
Manager, People and Society – World Business Council for Sustainable Development (WBCSD); Former Manager, Global Future Council on Media, Entertainment and Sport, World Economic Forum

Anne Josephine Flanagan
Project Lead, Data Policy, World Economic Forum

Eddan Katz
Former Platform Curator, Artificial Intelligence and Machine Learning, World Economic Forum

Sebastien A. Krier
Senior Technology Policy Researcher at the Stanford University Cyber Policy Center

Estelle Masse
Senior Policy Analyst and Global Data Protection Lead, Access Now

Emily Ratté
Project Specialist, Artificial Intelligence and Machine Learning; Manager, Global Future Council on AI for Humanity, World Economic Forum

Hesham Zafar
Community Curator, Media, Entertainment and Sport, Manager, Global Future Council on Media, Entertainment and Sport, World Economic Forum

Endnotes

1. Trauma is defined as an experience that produces psychological injury or pain. "Trauma-informed", within the context of this paper, refers to the creation of policies, regulations and redress mechanisms that keep this notion in mind in order not to cause further trauma to a victim of technology-enabled harms.
2. Data-driven harms result from the adverse effects caused by uses of data that may impair, injure or set back a person, group, entity or society's interests as a whole. Data-driven harms also include harms created by both intentional and unintentional uses of AI and predictive technologies. This term is used throughout the paper, often interchangeably with the term "digital harm".
3. A victim is a person who is targeted by some form of harm or abuse.
4. Autonomous sensory meridian response (ASMR) is a perceptual sensory phenomenon, likened to meditation, that encompasses a pleasant and calming "tingling" sensation localized to the scalp and neck.
5. Compton, Sophie, "More and More Women Are Facing the Scary Reality of Deepfakes", Vogue, 16 March 2021: https://www.vogue.com/article/scary-reality-of-deepfakes-online-abuse; #MyImageMyChoice is a coalition of survivors and advocates calling for a new bill to reform England's criminal law and government policy. The campaign is currently focused on legislation in England and Wales and will expand to focus on US legislation in early 2021. #MyImageMyChoice wants to overturn the fact that victims of image-based sexual abuse are routinely doubted and ignored. The project was set up by activist film-makers Sophie Compton, Reuben Hamlyn and Elizabeth Woodward, with the aim of creating a safe and supportive environment for victims to speak about their experiences – either privately or publicly. With consent, these testimonies are presented to lawmakers and those with the power to make change, ensuring that they have evidence they need.
6. "The State of Deepfakes 2019 Landscape, Threats, and Impact", *Sensity.ai,* 2019.
7. A deepfake is a digitally manipulated video that uses AI technology to swap the face of one person onto another person's body, usually without their consent. Dickson, E. J., "TikTok Stars Are Being Turned into Deepfake Porn Without Their Consent", Rolling Stone, 26 October 2020: https://www.rollingstone.com/culture/culture-features/tiktok-creators-deepfake-pornography-discord-pornhub-1078859/.
8. "The Ripple Effect: COVID-19 and the Epidemic of Online Abuse", Glitch UK and End Violence Against Women Coalition: https://www.endviolenceagainstwomen.org.uk/wp-content/uploads/Glitch-and-EVAW-The-Ripple-Effect-Online-abuse-during-COVID-19-Sept-2020.pdf.
9. Digital humanity refers to the quality or state of being human in the digital realm.
10. Ross, Alec, *The Industries of the Future,* Simon & Schuster, 2017: https://www.amazon.com/Industries-Future-Alec-Ross/dp/1476753660.
11. According to the GDPR Article 4, "personal data" means any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
12. United Nations, "Data Economy: Radical Transformation or Dystopia?", *Frontier Technology Quarterly*, January 2019: https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/FTQ_1_Jan_2019.pdf.
13. Defamation is a statement that injures a third party's reputation. In common law countries, tort of defamation includes both libel (written statements) and slander (spoken statements); Legal Information Institute, "Defamation", Cornell School: https://www.law.cornell.edu/wex/defamation.
14. Mental suffering is an emotional response to an experience that arises from the effect or memory of a particular event, occurrence, pattern of events or condition. Emotional distress can usually be discerned from its symptoms (e.g. anxiety, depression, loss of ability to perform tasks or physical illness). In common law countries, there are two causes of action that involve infliction of emotional distress: intentional infliction of emotional distress and negligent infliction of emotional distress – i.e. bystander action; Legal Information Institute, "Emotional Distress", Cornell School: https://www.law.cornell.edu/wex/emotional_distress.
15. "NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software", NIST, 19 December 2019: https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software.
16. Castelvecchi, Davide, "Is Facial Recognition Too Biased to Be Let Loose?", *Nature*, 18 November 2020: https://www.nature.com/articles/d41586-020-03186-4.
17. Directorate for Science, Technology and Industry, "At a Crossroads: 'Personhood' and Digital Identity in the Information Society", Organisation for Economic Co-operation and Development (OECD), 29 February 2008.
18. The OECD defines the term "person" as a human being or a natural person.
19. Wang, Tricia, "You Are Not Your Data But Your Data Is Still You", Deep Dives, 7 August 2020: https://deepdives.in/you-are-not-your-data-but-your-data-is-still-you-b41d2478ece2.
20. Reventlow, Nani Jansen, "Digital Rights Are *All* Human Rights, Not Just Civil And Political", Berkman Klein Center for Internet and Society at Harvard University, 27 February 2019: https://medium.com/berkman-klein-center/digital-rights-are-all-human-rights-not-just-civil-and-political-daf1f1713f7a.

21. Cook, Jesselyn, "A Powerful New Deepfake Tool Has Digitally Undressed Thousands of Women", Huffpost, 10 August 2021: https://www.huffpost.com/entry/deepfake-tool-nudify-women_n_6112d765e4b005ed49053822.
22. So far, only Facebook has banned the website's URL from its platform, in an effort to stop traffic and further proliferation of that content.
23. Non-consensual pornography (NCP) is defined as the distribution of sexually graphic images of individuals without their consent (Cyber Civil Rights Initiative).
24. Ibid.
25. The premise of foresight is that the future is still in the making and can be actively influenced or even created, rather than what has already been decided, there only to unearth and discover, and passively accepted as a given. This is an empowering realization for both governments and citizens. Foresight permits governments and public administrations to construct contingency plans for undesirable but possible and probable scenarios, while creating policies that capitalize on the transformational possibilities of preferred futures, moving from foresight and insight to strategy and action; Global Centre for Public Service Excellence, "Foresight: The Manual", UNDP, November 2014.
26. Redish, M. H. and Marshall, L. C., "Adjudicatory Independence and the Values of Procedural Due Process", *Yale Law Journal 95*, 1985: 455.
27. Miller, David, "Justice", The Stanford Encyclopedia of Philosophy, Fall 2017 Edition, Edward N. Zalta (ed.): https://plato.stanford.edu/archives/fall2017/entries/justice/.
28. Lesbian, gay, bisexual, transgender, queer and intersex.
29. Leufer, Daniel, "Computers Are Binary, People Are Not: How AI Systems Undermine LGBTQ Identity", *Access Now,* 6 April 2021: https://www.accessnow.org/how-ai-systems-undermine-lgbtq-identity/.
30. Ibid.
31. Fish, B., Bashardoust, A., Boyd, D., Friedler, S., Scheidegger, C. and Venkatasubramanian, S., "Gaps in Information Access in Social Networks?", The World Wide Web Conference, May 2019, pp. 480–490.
32. Coeckelbergh, M., "Artificial Intelligence, Responsibility Attribution and a Relational Justification of Explainability", *Science and Engineering Ethics* 26(4), 2020: 2051–2068.
33. Burrell, J., "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms", *Big Data & Society* 3(1), 2016: 2053951715622512.
34. Lipton, Z. C., "The Mythos of Model Interpretability: In Machine Learning, the Concept of Interpretability Is Both Important and Slippery", *Queue* 16(3), 2018: 31–57.
35. Lakkaraju, H. and Bastani, O., "How Do I Fool You? Manipulating User Trust via Misleading Black Box Explanations", *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, February 2020, pp. 79–85.
36. Siddique, Haroon and Quinn, Ben, "Court Clears 39 Post Office Operators Convicted Due to 'Corrupt Data'", *The Guardian*, 23 April 2021: https://www.theguardian.com/uk-news/2021/apr/23/court-clears-39-post-office-staff-convicted-due-to-corrupt-data. See Hamilton vs. Post Office Ltd, [2021] EWCA Crim 577 (23 April 2021).
37. Solomon, Brett, "Can Human Rights Survive the Digital Age? Only If We Do These Things", *Access Now*, 10 December 2020: https://www.accessnow.org/human-rights-in-the-digital-age/.
38. General Data Protection Regulation (GDPR), "Regulation [EU] 2016/679 of 27 April 2016": https://www.legislation.gov.uk/eur/2016/679/contents#:~:text=Regulation%20%28EU%29%202016%2F679%20of%20the%20European%20Parliament%20and,%28General%20Data%20Protection%20Regulation%29%20%28Text%20with%20EEA%20relevance%29v.
39. Cyber Civil Rights Initiative: https://www.cybercivilrights.org/ccri-press-releases/.
40. Department for Digital, Culture, Media and Sport, "Draft Online Safety Bill", GOV.UK, 12 May 2021: https://www.gov.uk/government/publications/draft-online-safety-bill. See also note 75.
41. Ibid.
42. The United Nations Human Rights Office of the High Commissioner, "International Covenant on Civil and Political Rights", 23 March 1976: https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx.
43. Martin, Noelle, "Only a Global Response Can Tackle the Rise of Online Harms. Here's Why", World Economic Forum Agenda, 5 August 2021: https://www.weforum.org/agenda/2021/08/only-global-response-tackle-rise-online-harms/.
44. Solomon, Brett, "Can Human Rights Survive the Digital Age? Only If We Do These Things", Access Now, 10 December 2020: https://www.accessnow.org/human-rights-in-the-digital-age/.
45. See note 57: http://www.npc.gov.cn/npc/c30834/202108/a8c4e3672c74491a80b53a172bb753fe.shtml.
46. Ibid.
47. E.g. New Zealand's Harmful Digital Communications Act 2015.
48. E.g. the UK's Malicious Communications Act 1988 or section 127 of the Communications Act 2003.
49. E.g. section 162 of Canada's Criminal Code (R.S.C., 1985, c. C–46) (as inserted in 2014); Ireland's Harassment, Harmful Communications and Related Offences Act 2020; in the US, 48 states (all but Massachusetts and North Dakota), the District of Columbia and Guam have enacted laws against non-consensual pornography (see Cyber Civil Rights Initiative: https://www.cybercivilrights.org/revenge-porn-laws/).

50. See, in particular, the EU (General Data Protection Regulation [GDPR] [Regulation {EU}] 2016/679 of 27 April 2016); this has been implemented by appropriate legislation in the EU's 27 member states. Similar legislation has been enacted in Brazil (Lei Geral de Proteção de Dados Pessoais [LGPD]); China (Personal Information Protection Law [PIPL], see note 52); Japan (Act on the Protection of Personal Information [APPI]); Kenya (Data Protection Act [DPA]); South Africa (the Protection of Personal Information Act [POPIA]); and the United Kingdom (Data Protection Act 2018). Similar legislation is pending in Canada (Bill C-11, to enact a Consumer Privacy Protection Act [CPPA] and a Personal Information and Data Protection Tribunal Act [PIDPTA], to update the Personal Information Protection and Electronic Documents Act [PIPEDA]); and India (Personal Data Protection Bill). In the US, omnibus privacy laws have been enacted in California (the California Consumer Privacy Act [CCPA]) and the California Privacy Rights Act [CPRA]); Colorado (the Colorado Privacy Act [CPA]); and Virginia (the Consumer Data Protection Act [CDPA]), with similar legislation pending in several other states. In July 2021, the Uniform Law Commission approved a Uniform Personal Data Protection Act (UPDPA) to provide a template for uniform state privacy legislation; the possibility of federal legislation is being explored in Congress (see https://iapp.org/resources/article/us-state-privacy-legislation-tracker/).

51. E.g. section 230 of the US Communications Decency Act (47 U.S.C. § 230); section 512 of the US Digital Millennium Copyright Act (17 U.S.C. § 512).

52. E.g. Articles 12–15 of the EU eCommerce Directive (Directive 2000/31/EC of 8 June 2000).

53. E.g. the US Fight Online Sex Trafficking Act (FOSTA) and Stop Enabling Sex Traffickers Act (SESTA) 2018.

54. E.g. the EU's proposed Digital Services Act (COM [2020] 825 final; 15 December 2020).

55. E.g. the EU's Digital Single Market Directive (Directive [EU] 2019/790 of 17 April 2019).

56. E.g. the EU's proposed Regulation on Terrorist Content Online (COM/2018/640 final; 12 September 2018).

57. Bills currently before the US Congress would – variously – significantly limit, reform or abolish the safe harbours set out above; they are collected at the Section 230 Reform Legislative Tracker, available at: https://slate.com/technology/2021/03/section-230-reform-legislative-tracker.html.

58. See Loi n° 2021–1109 du 24 Août 2021 LOI n° 2021–1109 du 24 Août 2021 Confortant le Respect des Principes de la République.

59. See the Netzwerkdurchsetzungsgesetz 2017 (NetzDG), as amended by the Gesetzes zur Änderung des Netzwerkdurchsetzungsgesetzes 2021, and as supplemented by the Gesetz zur Bekämpfung des Rechtsextremismus und der Hasskriminalität 2021.

60. See the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 [hereafter: the IT Rules] (see also note 67).

61. See Federal Law No. 482–FZ on Amendments to the Federal Law on Enforcement Actions Regarding Persons Involved in Violations of Fundamental Human Rights and Freedoms and the Rights and Freedoms of the Russian Federation Citizens (30 December 2020).

62. See Bill C-10 (proposed 3 November 2020); see also the Technical Paper on Addressing Harmful Content Online (published 29 July 2021); this whole issue became a prominent issue in the August 2021 federal parliamentary election.

63. See the Online Safety and Media Regulation Bill (published on 10 January 2020).

64. See the Online Safety Bill (the most recent version was published on 12 May 2021).

65. E.g. the proposed Digital Services Act.

66. Established by the Enhancing Online Safety Act 2015 (Cth); see also the Online Safety Act 2021 (Cth).

67. Reno vs. American Civil Liberties Union, 521 U.S. 844 (1997).

68. Shreya Singhal vs. Union of India, AIR 2015 SC 1523. The IT Rules (see note 72) are facing a similar challenge; see Agij Promotion of Nineteenonea Media Pvt. Ltd. vs. Union of India; and Nikhil Mangesh Wagle vs. Union of India (9 August 2021; interim orders).

69. See the decisions of the Conseil Constitutionnel (the Constitutional Council) in Décision n° 2020–801 DC du 18 Juin 2020 and Décision n° 2021–823 DC du 13 Août 2021. In the 2021 decision, the council struck down elements of the Projet de loi: Respect des Principes de la République; the remainder of that bill came into force as Loi n° 2021–1109 du 24 Août 2021.

70. See the decision of the Bundesverfassungsgericht (the Federal Constitutional Court) in 1 BvR 1873/13, 1 BvR 2618/13 (27 May 2020) (*Bestandsdatenauskunft II; Subscriber data II)*. Similarly, Google and its subsidiary YouTube have challenged various elements of the 2021 legislation in the Cologne Verwaltungsgericht (Administrative Court) as contrary to the Grundgesetz (the German Constitution) and EU law; see Frank, Sabine, "Zum Erweiterten Netzwerkdurchsetzungsgesetz in Deutschland – Anmerkungen von YouTube", YouTube Official Blog, 27 July 2021: https://blog.youtube/intl/de-de/news-and-events/zum-erweiterten-netzwerkdurchsetzungsgesetz-deutschland/. An important case, seeking similar procedural safeguards for Facebook's content moderation practices, is ongoing in Poland (see Społecznej Inicjatywy Narkopolityki vs. Facebook Ireland Limited Ref. IV Th 97/20 p-I: 14 May 2021; interim orders).

71. Application no. 20159/15, Bulgakov vs. Russia (23 June 2020); Application no. 61919/16, Engels vs. Russia (23 June 2020); Application no. 10795/14, Kharitonov vs. Russia (23 June 2020); Application nos. 12468/15, 23489/15 and 19074/16, OOO Flavus and others vs. Russia (23 June 2020). See also Application nos. 58170/13, 62322/14 and 24960/15, Big Brother Watch and Others vs. the United Kingdom (25 May 2021).

72. For example, FOSTA and SESTA were criticized not only by civil liberties organizations but also by representatives of consensual sex workers, who argued that the bills drove sex work further underground and thus made it more dangerous. See: https://medium.com/@jmalcolm/fosta-sesta-isnt-just-an-attack-on-sex-workers-it-s-also-an-attack-on-free-speech-f764f9c09452. It led to consequences far beyond those limited to its intended target of sex trafficking, including the closure of Craigslist's entire Personals section: https://www.craigslist.org/about/FOSTA.

73. Angwin, Julia, Larson, Jeff, Mattu, Surya and Kirchner, Lauren, "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks", ProPublica, 23 May 2016: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

74. Choholas-Wood, Alex, "Understanding Risk Assessment Instruments in Criminal Justice", The Brookings Institute, 19 June 2020: https://www.brookings.edu/research/understanding-risk-assessment-instruments-in-criminal-justice/.

75. McDonald, Sean, "Our (Mis)represented Digital Rights", Center for International Governance Innovation (CIGI), 26 May 2021: https://www.cigionline.org/articles/our-misrepresented-digital-rights/.

76. Danielle Citron describes "technological due process" as the extension of historic doctrines and protections of procedural due process into the public and government processes that are increasingly modelled and administered through technology. While every legal culture has its own framework for procedural fairness, there are comparatively few digital and technological articulations and implementations of those rights. See Citron, Danielle Keats, "Technological Due Process", Washington University Law Review 85(6), 2008: https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview.

77. The World Bank, "Unraveling Data's Gordian Knot: Enablers and Safeguards for Trusted Data Sharing in the New Economy", Washington, DC, 2020: https://documents1.worldbank.org/curated/en/863831612427670947/pdf/Unraveling-Data-s-Gordian-Knot-Enablers-and-Safeguards-for-Trusted-Data-Sharing-in-the-New-Economy.pdf.

78. BIPOC stands for: Black, Indigenous and/or people of colour.

WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744
contact@weforum.org
www.weforum.org