The case for collective action against the harms of data-driven technologies

To what extent are the GDPR's data rights an effective tool for enabling collective action?

Ausloos, J.; Toh, J.; Giannopoulou, A.

Ada Lovelace Institute

The case for collective action against the harms of data-driven technologies

To what extent are the GDPR's data rights an effective tool for enabling collective action?

Jef Ausloos , Jill Toh , Alexandra Giannopoulou

23 November 2022

Reading time: 20 minutes

This is the first of three blogs in a series that explores the potential of collective action, within and beyond existing legal mechanisms, as a means of redress against the harms of data-driven technologies.

Introduction

While the implementation of data-intensive technologies – such as automated decision-making systems, surveillance tools, facial recognition technologies and predictive analytics – may provide certain benefits, there is clear evidence that they perpetuate and reinforce existing power asymmetries.

Vast amounts of data, and the technological means to leverage it, are often in the hands of already powerful actors, whether private entities or public authorities and generate new risks and harms at a speed and scale previously unattainable. Loss of privacy, discrimination, inequality and exploitation are among the harms and injustices that have emerged or have been exacerbated through the deployment of data-driven systems.

For instance, data analytics support algorithmic decision-making that sorts and categorises individuals on the basis of distinctive characteristics, leading to various kinds of

inequalities. Among them, intersectional forms of discrimination are becoming increasingly complex and difficult to prove and address, as evidenced in current anti-discrimination law discourse. Because they are mediated through algorithmic systems, these harmful decisions tend to be coated in a veneer of neutrality.

These data-driven harms occur across the intersecting and often-overlapping levels of the individual, the collective and society. In this three-part blog series, we focus on the *collective* dimension of data-driven harms, which has so far received less attention than the individual and society-wide ones, and on *collective action* as a means to access rights. As a swathe of EU legal proposals are being tabled we look at one of the pieces of legislation that kickstarted the present era of technology law, the GDPR. We evaluate the extent to which it recognises collective harms and whether its application has enabled collective action.

People are often not fully aware of the extent to which their environment is powered by algorithmic systems and how it may be harmful. As a result, data-driven processes widen the gap between people and the decisions or events that affect their lives while obfuscating forms of collective discrimination, i.e. harms that affect specific groups of people. This opacity and disconnect, caused by technical issues and socio-legal barriers, result in concrete obstacles for those that intend to challenge an outcome or the processes that led to it.

For this reason, embracing collective action to confront data-related harms is a strategy that deserves particular attention, especially in contexts or scenarios that are already structured by inherently asymmetric power dynamics, such as those between employer and employee, migrants and states or students and institutions.

In this first instalment, we foreground concrete instances of the collective impact of data-driven technologies in the contexts of work, migration and education, and explain how harms are exacerbated by existing power and information asymmetries. We then consider the role that present regulation, and especially GDPR data rights, can play in supporting collective action.

Using ‘collective action’ to confront data harms

The legal understanding of collective action is varied and unclear. In legal terms, collective

action is usually understood as the right to procedural class action and/or to the positive protection of a collective interest, in which (data) rights can be utilised. This includes countering the negative effects of data-driven technologies, for instance, in strategic litigation or in a court case.

However, collective action extends beyond legal action *ex post* (meaning action that happens after the fact, through litigation and other means of accessing justice) and legal protection *ex ante* (meaning through preventive measures). In this blog series, we use the term beyond its strict legal understanding, to mean both the situations in which a number of individuals coordinate their actions (i.e. to collectivise) *and* the development of a 'collective agency'[1] pursuing a goal that would benefit individual subjects as well as the collective they participate in as a whole.

Notably, given the opaque nature of data-driven technologies, affected individuals may not know about each other. The identification of a group of people that is being harmed as a collective is a contextual exercise, as groups tend to be in flux, overlapping or formed through externally imposed factors.

As we will see later in this series, alongside other institutional efforts, GDPR's data rights, such as the right of data access – which gives individuals the right to obtain their personal data from a public or private organisation that holds it – can be deployed to understand otherwise opaque algorithmic systems and data infrastructures, identify underlying characteristics (such as profession or migration status) by which data-driven technologies systematically harm groups of individuals and build the awareness necessary to foster collective agency.

In other words, data rights are a pragmatic necessity for progressive action and movements in digital environments, as they can be used to collectivise groups of people. Strength in numbers is key when countering the negative effects of data-driven technologies. While no silver bullet, data rights can help bring together groups of people who share related issues – despite the shifting, overlapping nature of group identities – and place agency in their hands to build momentum for broad sets of collective goals.

Three contexts of collective data-driven harms

The examples below highlight how individuals can be harmed through data-driven

technologies *because* they belong to a specific population (i.e. asylum seekers living in camps) or because of their occupation (gig workers and students), and so suffer the consequences of pre-existing power asymmetries.

Many people argue that the operators of algorithmic systems, whether private or public actors, do not *intend* to cause harm. However, the effects of their actions, which may end up being discriminatory, exclusionary or simply unfair, are neither abstract nor theoretical and impact people's livelihoods in a myriad of ways.

While the harms identified in the examples evidently affect individuals, they can be said to experience harm collectively. This is because these individuals belong to a group of people or a population that can be externally identified either through technology (i.e. extractive data processes or algorithmic governance, categorisation and sorting) or by law (i.e. legal categories).

Recognising this means that we can extend the claim that a single person is being harmed to all individuals belonging to, or constituting, the specific group, collective or population that the individual is part of. While the law has characteristically had trouble understanding and incorporating the logics of collective harm directly, it does – through procedural processes – already enable groups of subjects to challenge these systems indirectly.

Algorithmic management and surveillance of on-demand delivery workers and drivers

The platform economy has been a testbed for workplace surveillance technologies, where companies such as Uber, Glovo, Deliveroo and Bolt practice widespread and continuous data collection and processing, and implement algorithmic management at scale. This has created significant harms and negative implications to large numbers of platform workers, in Europe and globally.

The way labour is organised via processes of datafication and forms of algorithmic management gives employers unprecedented granular control and surveillance over the entire workforce. Workers' schedules, wages, rewards and ranking systems through gamification strategies, for instance, are tightly controlled and mediated by data and algorithmic systems.

Workers are ‘deactivated’, that is dismissed through automated or semi-automated means, due to alleged ‘fraudulent activity’ and breaching of contractual terms that are unilaterally set and tweaked by the companies. As of 2019, there were more than 1,000 (known) cases in the UK of alleged fraudulent activity, which have resulted in terminations.

The use of discriminatory facial recognition systems for identity verification is one of the sources of collective harm on workers. For instance, Uber’s ‘Real Time ID Check system[2] – proven to produce high error rates when used on women and people of colour[3] – requires workers to conduct selfie identity checks when logging on to the app for verification purposes. In several known cases, the system’s inability to accurately recognise faces has led to misidentification or a failure of these checks, resulting in workers being unable to operate the app (and, in some cases, with any other company that uses the same license).

In the USA, where data-sharing laws are more relaxed than in the EU, ride-hailing and delivery platform companies have started outsourcing background checks to third parties, as well as creating joint databases on alleged workers who have been accused of assault.

Operating in these conditions, whereby the risks – such as pressures on delivery target times or fear of terminations – are entirely offloaded on the workforce, significantly compromises workers’ psychological and physical health and safety.

While many of these incidents, at first glance, are harms inflicted on individuals, the number of cases of collective harm, evidenced by the growing number of court cases on employment, discrimination and data protection infractions, brought forward by workers, unions, activists, lawyers, researchers and civil society, are cause for significant concern.

Simultaneously, many of the relatively recent regulatory improvements in the context of the platform economy have been achieved through organising, mobilisation and campaigning

towards collective actions.

AI surveillance tools for migrants in refugee camps in Greece

Centaur, the partly automated surveillance system funded by the European Union, is being deployed at camps for refugees and asylum seekers in Greece. Monitoring with this new data-powered system means that the camps operate as 'highly surveilled prisons', incorporating a variety of tools such as CCTV systems, drones, perimeter violation alarms with cameras, control gates with metal detectors and X-ray devices, fingerprint control on entrance and exit, and an automated system for broadcasting public announcements from loudspeakers.

These tools are deployed to maximise state control on migrant populations and are causing concern among the residents of the camps and civil society. Viewed in the context of an increasingly racist and xenophobic environment that has exacerbated the exclusion and criminalisation of refugees, the EU is violating the basic freedoms of vulnerable individuals and populations by datafying their bodies, movements and relations as inputs for an automated management system.

Centaur is one artefact in the broad range of automated tools applied in border control and migration management. As pointed out by EDRi, the extractive data processes and opaque nature of the decision-making features of these tools 'create an environment ripe for algorithmic discrimination', further aggravating the already precarious lives and vulnerability of migrant populations.

Notably, the Greek Government has confirmed that Centaur is GDPR compliant,[4, 5] but without offering any evidence to support its statement. But even if Centaur was 'GDPR compliant', this is unlikely to mitigate its negative effects and injustices.

Indeed, GDPR compliance in this case does not necessarily translate into the capacity of single individuals to exercise their data rights without coordination. There are many

pressing practical and systemic obstacles – characterised by various forms of power and information asymmetries, including language access – that migrants have to overcome, if they are to claim their data rights. And data rights may not be a high priority for people who already face severe hardship.

This power dynamic compounds the existing effects of *a priori* established imbalances between border control forces and national governments, on the one side, and migrants, refugees and asylum seekers, on the other. In practice, there remain limited (or no) possibilities for an individual migrant, refugee or asylum seeker to exercise their rights.

The role that collective protection can play against data infrastructures such as Centaur is crucial and should be emphasised over the restrictively interpreted 'GDPR compliance' that the data-intensive system supposedly fulfils.

Proctoring software and students' personal data

The lockdown measures imposed from the beginning of 2020 as a response to the COVID-19 pandemic meant that universities all over the world quickly migrated online. This rapid shift towards the provision of virtual education has been characterised by the externalisation to and use of third-party service providers, such as Zoom, to ensure continuity.

Among the tools that stand out in the 'suite' of data-driven educational services is e-proctoring software. With the risk that pandemic-related lockdowns would negatively affect exams, educators and educational institutions focused on recreating in-classroom invigilation online, through specialised software.

Many schools decided to rely on e-proctoring systems, which allow the (a)synchronous monitoring of students by taking over a computer system's webcam and microphone. E-proctoring is enabled by a set of machine learning tools operating as an algorithmic watchdog that checks hundreds of students in real-time and gives each of them a score,

which is supposed to represent how ‘aberrant’ their behaviour was during an exam.

In addition to issues related to intrusion of privacy and contributing to the stress of taking exams, e-proctoring software has repeatedly shown signs of discrimination and racism. There is reliable evidence that it is unable to detect darker skin tones and disproportionately penalises students of colour. The use of such kind of software is exclusionary also in other ways, including the assumption that all students have access to stable internet connection and learning-compatible home environments.

The reliance of e-proctoring systems on surveillance capabilities has resulted in students’ protests, including resisting the collection, use and processing of their personal (and sensitive) data, and litigation.[6, 7, 8] However, countering the effects of this type of software is proving difficult.

While individual harms are evidently occurring, it has also become clear that any challenge to the use of e-proctoring software will require coordination and collaboration in order to succeed. Recognising that the software risks negatively affecting students and, more broadly, that its use relies on techno-solutionist arguments to boost a pedagogy of punishment will hopefully lead to more comprehensive and collective action against it.
In conclusion, there are few incentives, expected advantages or guarantees for an individual to exercise their data rights, particularly when asymmetrical power dynamics are at play. It is exactly for this reason, that we need to emphasise and value the role of collective protection from, and action against, data infrastructures and algorithmic systems. *Ex-ante* protection should account for the significant power asymmetries structuring the relationships between those deploying data-intensive technologies and those affected by them, and *ex-post* responses should include collective actions coordinated by NGOs or other civil society groups.
An enabling environment for collective action will have the added benefit of mitigating the legitimate hesitance by marginalised populations to take action, due to fears concerning their position and the pressing survival problems they are likely to face.

Data rights setting the basis for collective action

The above cases paint a bleak picture of how both the law and technology are complicit in producing, sustaining and exacerbating data-driven (collective) harms. They also show that, without a tool to identify the characteristics shared by those who suffer harm, it is harder to reveal how data-driven technologies affect specific groups of people and to shape appropriate legal responses. This is where the already mentioned data rights from Chapter V of the GDPR offer some limited promise and incentivise the kind of collective action that is necessary.

The opportunity for collective action within the context of data subject rights and the GDPR

Due to the wide scope of application of the GDPR, the rights (and obligations) it includes are quite abstract. Data subject rights are purpose-blind or intent-agnostic and are formulated so that they can be invoked in an infinite number of situations. Indeed, data subject rights can be used by anyone whose fundamental rights, freedoms or interests are affected by the processing of personal data. As such, they hold particular promise to safeguarding diverse groups of individuals and communities – from platform workers, to refugees or students – at least in theory.

For instance, the right of access under the GDPR has already been used strategically by different interest groups to break the information and power asymmetries affecting them, although with mixed results. The exercise of this right can (and has been) instrumental not only in concretely identifying the violations that constitute the legal enabler for litigation, but, more importantly, in empowering people to collectivise, by helping them recognise the ways in which they are affected, not only as single individuals, but as groups.

These ‘alternative’ pathways of empowerment through collective action, mediated by GDPR data rights, could support whole groups of individuals against systemic threats and the entrenched power asymmetries imposed by state institutions and companies.

The limitations of data subject rights for collective action

While enabling their polyvalence, however, the abstract phrasing and wide scope of data subject rights have considerable downsides, as they essentially push interpretation costs downstream to the parties invoking and accommodating data subject rights. In fact, this

aspect of data rights can further solidify existing power asymmetries, giving well-resourced entities significant interpretational leeway and few concrete red lines for weak(er) parties to fall back on.

Moreover, despite the GDPR's explicit aim to make data subject rights as easily invokable as possible, they are often expressed in terms that are too far removed from the practical realities in which they can be most valuable. Indeed, individuals in vulnerable positions, who are affected the most by data processing infrastructures and their harms, may not even be aware of their data subject rights in the first place. And, as illustrated through the examples above, the focus on (and tendency to load responsibility on) the individual, benefits powerful data-processing organisations and actively harms people.

Data protection rights have long been associated with the notion of individual privacy. Indeed, they have often been characterised as a subset of 'informational' privacy protection. Yet, this understanding is quite narrow, and fails to consider the history, scope and rationale of data protection law.

However, while data protection has never exclusively focused on the protection of individuals, this has certainly become the dominant narrative over the years. The exact reasons for this are hard to discern, but arguably include the adoption of the Charter with its right to 'protection of personal data' and, more generally, the neoliberal turn in policymaking, with a strong emphasis on individual rights, liberties and responsibilities.

In our next blog post on collective action and data rights, we will consider how power asymmetries have affected and affect all phases of the GDPR, from inception to enforcement, reproducing the emphasis on individual rights and limiting the possibilities for collective action.

This blog series is published in the context of our Rethinking Data research project, which sets an ambitious vision for data use and regulation that can deliver a positive shift in the digital ecosystem towards people and society. Read our report here.

Project

Rethinking data and rebalancing digital power

Keywords

AI and data ethics

Data governance

Data regulation

Europe

Authors

Jef Ausloos

Jill Toh

Alexandra Giannopoulou

Related content

Blog

How the GDPR can exacerbate power asymmetries and collective data harms

Jef Ausloos , Jill Toh , Alexandra Giannopoulou

Exploring how power asymmetries operate across the law and collective harms

29 November 2022

AI and data ethics Data governance ...

Report

Rethinking data and rebalancing digital power

Valentina Pavel

What is a more ambitious vision for data use and regulation that can deliver a positive shift in the digital ecosystem towards people and society?

17 November 2022

Data regulation Digital inequality ...

Completed project

Rethinking data and rebalancing digital power

What is a more ambitious vision for data use and regulation that can deliver a positive shift in the digital ecosystem towards people and society?

Data regulation Society, Justice & Public Services

Blog

The role of collective action in ensuring data justice

Jef Ausloos , Jill Toh , Alexandra Giannopoulou

Five preconditions to protecting people from data-driven collective harms

1 December 2022

AI and data ethics Data governance ...