Individual property rights or collective democratic governance? Privacy in the age of AI

Maximilian Kasy

Why should we care about privacy, and how should we define it? The answer depends on whose perspective we adopt and whose interests we prioritise. We might take the perspective of a data collector, such as a technology company collecting user data or a state agency conducting a census. Alternatively, we might adopt the perspective of the data subjects themselves, such as the users of a digital platform or the citizens of a country. These perspectives lead to different recommendations - individual privacy protections versus collective data governance - because of data externalities. Data externalities are ubiquitous in the age of machine learning and artificial intelligence (AI).

Data collectors and data subjects

From the perspective of data collectors, the primary goal is to eliminate incentives for data subjects to withhold their data. If data subjects can be sure that they will not face any adverse consequences by sharing their individual data, they will be more willing to do so. This goal can be achieved through differentially private mechanisms. Such mechanisms are designed so that an individual's decision to share data has, with very high probability, no consequences for that individual themselves.

Data subjects, as a group, have a different goal. They are concerned with ensuring that the collection and analysis of their collective data have no adverse downstream consequences. In general, neither individual privacy rights nor differentially private mechanisms can guarantee such collective protection from adverse consequences.

Data externalities

The divergence between the interests of data collectors and those of data subjects arises from data externalities. While a single person's data sharing may have negligible consequences for themselves, the aggregate effect of everyone sharing their data may have significant and harmful consequences for the group as a whole. This dynamic parallels carbon emissions: the impact on any individual emitter of their own emissions is negligible, yet the global consequences may be catastrophic.

Data externalities are intrinsic to statistical learning, and to the extraction of patterns across individuals. Data externalities do *not* arise in settings where individual-level decisions are made by fixed algorithms or bureaucratic procedures that operate solely on individual data. Data externalities have, however, become pervasive in the age of machine learning and artificial intelligence. Machine learning is fundamentally concerned with identifying patterns across individuals rather than evaluating isolated data points. By sharing their data, individuals enable learning algorithms to make predictions about people similar to themselves; predictions that may inform consequential decisions affecting those similar people.

When data externalities are present, the introduction of differential privacy can make the data subjects collectively worse off: By eliminating individual incentives to withhold data, differential privacy facilitates more extensive data collection, which in turn enables algorithms to infer patterns that may be used to harm data subjects in various ways. In this respect, differential privacy can be analogous to a hypothetical subsidy of carbon emissions, implemented in an attempt to compensate individuals for the harms of climate change - which would be a rather questionable policy.

Examples

The abstract logic sketched above applies to many different settings of real-world importance, where the data collectors might be either private corporations or state entities such as police or military forces: Information filtering on social media platforms such as Facebook, the setting of individual wages and provision of work opportunities by gig work platforms such as Uber, the setting of individually tailored prices on shopping platforms, the automated selection of job-candidates for interviews, the targeting of individuals for incarceration and deportation by Immigrations and Customs Enforcement (using *ImmigrationOS*, provided by the company Palantir), or the selection of individuals to be bombed in Gaza (using systems such as *Lavender* and *WhereIsDaddy*).

To make things concrete, take the following (hypothetical) example: Consider a population of job seekers. Some of these job-seekers will have care-responsibilities in the future; they will have children, or they might have older family members who require their care. Employers have an interest to filter out candidates with such care responsibilities. Knowing this, the job candidates will be reluctant to reveal any information about their future care obligations.
Suppose now that a differentially private mechanism is in place to collect information about applicants' plans to have children etc. Because of differential privacy, individuals run no risk by disclosing this information. But once everyone discloses, machine learning allows the data collector to form predictions for each individual, to predict how likely they are to have

children, etc. With rich enough data and good enough models, it will be possible to predict these outcomes quite well. This then allows the employer to exclude future parents from consideration. Collectively, the job seekers would thus have been better off by withholding their data, even though individually there was no reason to do so.

To describe this logic in more technical terms, suppose that individuals i are characterised by a sensitive outcome Y_i (for instance, whether they will have care responsibilities), as well as by range of publicly available features X_i. Differential privacy guarantees that data subjects can share their Y_i without any fear of personal repercussions. But the goal of machine learning is not simply to learn the individual Y_i, the goal is to learn the relationship between Y_i and X_i; for instance the coefficients β of a predictive regression $\widehat{Y}_i = X_i \cdot \beta$. It is possible to learn these coefficients β, which describe patterns across individuals, while maintaining differential privacy for individual outcomes. If the features X_i are predictive for the outcome, so that $\widehat{Y}_i \approx Y_i$, then it is in turn possible to reconstruct each of the individual outcomes Y_i based on knowledge of β - despite differential privacy being maintained. If any one individual shares their Y_i, this improves predictions for all other (similar) individuals; this is the source of data externalities.

Collective control

In many cases, the interests of data subjects should be given normative priority over the interests of data collectors. Because individual privacy rights and differentially private mechanisms do not protect these interests, alternative forms of regulation and governance are required. Ultimately, we need collective democratic control by data subjects, to decide which data are collected, which patterns are inferred from these data, and what decision systems are built leveraging these patterns. Data subjects must have the right to determine which use cases are beneficial to them, and which are harmful.

Such collective control over data could be built on a case by case basis, for different application domains. This could be at the level of your local school board or municipal police department, your health care provider or employer, at the level of all workers on a gig platform or all users of a social media platform, etc. Collective democratic control would require that the stakeholders (data subjects) have a say over what data is collected, and to what ends. They need to decide on the objectives being maximised by any machine learning system trained on these data, in particular.

One attractive institutional arrangement for such democratic control is sortition (also known as citizens' councils): A random small set of individuals is selected to be representative of all the data subjects. The selected individuals are appropriately compensated, and given time and access to expertise, to learn about the data collected and the possible uses it

might be put to. They then make binding decisions about the actions and objectives of any algorithms using this data, to prevent harmful applications and promote desirable applications. Democratic control of this form, and not individual privacy rights or differentially private mechanisms, are the only way to make sure that AI is built and deployed in a way that serve all of us.