Position Statements on Responsible Uses of Technology and Health Data During Times of Crisis - Even desperate times call for “Fair Trade Data”

Magali Feys

Introduction

As a result of the COVID-19 crisis, demand for health data (and the importance of it) has rapidly increased. Tracking, understanding, and coming up with solutions to keep COVID-19 under control have led to a significant interest in the use of AI and algorithmic technologies to assist human actors. However, while AI can be extremely useful, it also carries a number of risks for the privacy of individuals, especially when dealing with health data. Now is the time to put safeguards in place to ensure that individuals are protected while still fostering an environment that encourages innovation, particularly in areas such as health that can provide huge benefits to society. Fair Trade Data standards can help to guide the application of AI, to drastically reduce the risk that AI causes unintentional harm to any group or individual.

Position

Realizing the full potential of AI could provide numerous and significant benefits to society. However, with increasing concerns over potential bias, discrimination and violations of data subjects’ privacy, these benefits could be lost. One critical thing to remember is that an algorithm or AI program is only as good as the data fed into it: incorrect data or poorly-designed algorithms can have serious and widespread consequences for individuals.

The idea of Fair Trade Data is that like with other resources, data should be handled and consumed in a way that is ethical, and does not harm individuals along the chain of use, whether at the point of production or the point of consumption. The technical safeguards embodied in Fair Trade Data are designed to maintain fidelity and reduce the possibility of re-identification, bias and discrimination while maintaining the highest levels of trust in the observations and decision-making resulting from its use. “Fair Trade Data” refers to data that has embedded, technically-enforced, granular privacy controls to eliminate the risk of “Conflict Data”: data that contains the risk of personal information being used to the disadvantage of that person. It is analogous to “conflict diamonds” being used against a country in which they are illegally mined to the disadvantage of the country.

To date, the opportunity presented by increasingly sophisticated data science technologies has been undermined by the widespread use of inadequate privacy enhancement techniques (PETs). Most PETs purport to protect the identity of the data subjects in any one dataset but fall far short of this goal in today’s big data world. The application of traditional PETs (including anonymization in combination with generalization and de-identification) introduces significant distortion into data. This may result in erroneous conclusions being drawn from protected datasets when compared to the original data in a non-protected form. The use of these technologies in the realm of AI contains serious risks when using health data in a widespread manner.

The use of Fair Trade Data is critical for creating much-needed transparency around the provenance of input datasets used to train AI applications. Data controllers and processors should be required to ensure that the steps they take in connection with AI processing limit the risk to individual data subjects of the misuse of their personal data against them. The rights of data subjects may be sacrificed going forward with no alternatives for recovery if companies (i) elect to fight in court rather than change the processes in which they have invested, and (ii) decide that the most cost-effective course of action is "regulatory arbitrage" against a perceived low risk/cost of enforcement action.

Current industry practices have outpaced the ability of policies alone or outdated technical approaches to adequately protect against bias, discrimination and violation of fundamental rights of privacy. Current data processing capabilities and practices require new Fair Trade Data principles that enforce:

- Data Protection by Design and by Default. Protections can be embedded into data to reduce the risk of re-identification. The use of technical and organisational safeguards such as GDPR-compliant Pseudonymisation, enables research institutions to authorise only specific uses of their data, with the flexibility to approve later further use while protecting the fundamental rights of the parties involved.
- Technical and organisational safeguards required for data and advanced processing to be legal: The GDPR requires technical and organisational safeguards that:
    - (a) transform non-compliant pre-GDPR data so that it remains legal to possess and process; and
    - (b) support a lawful basis for advanced analytics, AI, marketing, and other iterative processing applications to be conducted under the GDPR.
- Data use minimisation (vs collection or retention minimisation) by dynamically controlling re-identification: Maximise authorised and minimise unauthorised uses of data by dynamically reducing re-identification risks.
- Transparency and audit controls: Enable the availability of statistical properties of data sets to aid in interpreting decisions made using the data and to ensure auditable compliance with data privacy and use policies.
- Cross-sectional policy enforcement: Enable common data store(s) to programmatically support data protection and privacy rights management policies applicable to different entities and locations (i.e. companies, industries, states, countries, regions, etc.), and to do so simultaneously.
- Real-time policy adjustment: Adjust in real-time to changing requirements by modifying the intelligible form of data into which dynamically-obscured data are transformed.

These Fair Trade Data principles are consistent with the intentional omission of any "grandfather" provision under GDPR Recital 171 as well as the principles of lawfulness, purpose limitation, data minimization and data protection by design and by default under GDPR Articles, 5(1)(a), (b) and (c), and Article 25.

Summary

The COVID-19 crisis has emphasised how AI can be used to assist and support critical issues in society, but has also shed light on the potential risks of surveillance, privacy, and discrimination. Even when AI can benefit society for use in fighting pandemics such as COVID-19, approaching AI regulation with new frameworks of use is vital. Carefully selecting and using data to prevent bias in AI is fundamental to its ethical use: a Fair Trade Data approach presents an opportunity to maintain information fidelity while reducing the likelihood of re-identification, bias and discrimination against individuals, in line with established principles in the GDPR.