

Data Rules for Machine Learning: How Europe can Unlock the Potential While Mitigating the Risks

By Blanka Soulava, Hamish Cameron and Victoria Ying

Advised by Trey Herr and Bruce Schneier

Europe has a unique opportunity to set data rules that unlock the potential of machine learning while mitigating the risks.

Data rules are critically important to the way future societies will function, including:

- **An economy's capacity to take advantage of machine learning** – which will increasingly underpin innovation and international economic competitiveness.
- **Dictating who can surveil one's life and on what grounds** – which can undermine fundamental rights and exacerbate existing social inequalities.
- **Global digital norms** – which will shape how free, fair and inclusive societies are, how open they are to trade and investment, and how protected individuals are from harm.

The EU is lagging behind in machine learning. It is outmatched in terms of AI-related private investment, patent filings and data market growth. Spurred by great power competition, China and the United States are positioning their laws, bureaucratic structures, and public resources to further accelerate the deployment of AI technologies. Aware of the challenges, the European Commission has made “a Europe fit for the digital age” a priority to 2024. It is **pursuing a “third way” between China and the United States** characterized by a raft of legislative proposals that grapple with striking the right balance between unlocking the economic benefits while minimizing the risks to individuals and society.

The stakes are high. If Europe's AI trends continue, it will face higher costs to deploy machine learning. The resulting lag on native innovation could **shrink the EU's share of an increasingly digitized global economy** and **reinforce its dependency on foreign technologies**. Yet, if the EU implements short-term, innovation-driven policies, there is a high risk it embeds systems that **fail to protect fundamental rights and exacerbate social inequalities**. If it cannot get this balance right, the EU's inability to present an effective regulatory model will **undermine its ability to shape global digital rules and norms**.

The right data policies will be key if the EU is to succeed. This report provides European decision makers with 15 recommended actions designed to navigate the complex equities at stake based on close examination of the **geopolitical considerations** of the EU, the **domestic considerations** of EU member states, and the **commercial considerations** of EU-based businesses.

The recommendations aim to achieve the following goals:

1) Leverage the scale of Europe by addressing fragmented open data rules and procedures across Europe and harmonizing technical data standards to increase the quantity of data available for innovation and reduce the complexity and cost of preparing data for machine learning.

2) Balance economic autonomy and openness by promoting diversification in cloud services while maintaining competitive cloud offerings, simplifying rules to minimize legal ambiguity around data transfers, and removing anticompetitive barriers to data while incentivizing investment in data sets.

3) Protect fundamental rights by reducing the complexity and legal uncertainty of complying with data protection, strengthening enforcement capabilities, investing in machine learning techniques that reduce the need for centralized pools of personal data, and addressing gaps in antidiscrimination laws.