On the Rival Nature of Data Use in the Context of Privacy – Legal Applications

Ayelet Gordon-Tapiero, Dept. of Computer Science, Hebrew University;
Muhammad Saad, McCourt School of Public Policy, Georgetown University;
Katrina Ligett, Dept. of Computer Science, Hebrew University;
Kobbi Nissim, Dept. of Computer Science, Georgetown University

Data have become a cornerstone of the modern economy, driving innovation, informing decision-making, and reshaping societal interactions. It is important to understand the nature of data as an economic good to design an appropriate regulatory policy towards their allocation, management and use. Most of the existing scholarship views data as non-rival good, where use by one party does not restrict availability for use by others (Varian, 2018; Veldkamp & Chung, 2024).

In this Article, we argue that privacy considerations require us to treat the use of data as a rival good. That is, use of data by one party should impact use by others when considering privacy risks. While privacy protections enable the use of sensitive data, regulators need to carefully balance informative use and data subjects' privacy. One party's individual use of data may not violate privacy. However, unconstrained data use by many will lead to accumulation of privacy harms. Each time a privacy-sensitive dataset is queried, information about the underlying data is revealed, and a certain level of privacy is lost. We therefore advocate for regulatory mechanisms that control and limit data use in privacy-sensitive contexts and treat it as a rival good by limiting the number of queries answered, limiting queries' accuracy, or both.

Our research is a cross-disciplinary effort between computer scientists, legal scholars and public policy practitioners. We rely on information theory perspectives to draw a distinction between *data* as encoded information and *data use* as information-based releases as two separate economic goods. We argue that while data can be characterized as rival or non-rival, or excludable or non-excludable based on different use cases, data use should always be treated like a rival good in privacy-sensitive contexts. Using existing research on privacy risks (Dinur & Nissim, 2003; Gordon-Tapiero, Ligett, & Nissim, 2025), we establish that accumulation of data use can lead to leakage of sensitive information. The *Fundamental Law of Information Recovery* (Dwork & Roth, 2014) also shows that regardless of the use of any privacy preserving technique, unconstrained data use will eventually lead to a breach of privacy. While the phenomenon of privacy harms accumulating has been recognized in a limited number of legal contexts, including in the United States, we argue for its broader recognition and for new principles that can address this concern.

We therefore propose a set of regulatory principles that explicitly recognize the rival nature of data use. First, data use should be governed through systematic tracking and coordination across uses and users. Privacy harms accumulate across repeated uses. Effective governance therefore requires accounting for cumulative privacy loss and coordinating on competing data use claims for a shared privacy budget. Second, governance frameworks should call for an overall reduction in data use. This reduction can be done through limiting the number of information releases, constraining the amount of data accessed per use, prioritizing analyses that have lower privacy costs, and allowing for lower accuracy in queries, where possible. Third, we argue for the use of synthetic data as a complementary governance tool. When developed with privacy guarantees, synthetic data enable data reuse without depletion of the privacy budget. Taken together, these principles can enable the design of a policy framework that gives regulators the ability to balance data's informative use and protect data subjects' privacy.

We extend our principles to existing data protection regulations in the European Union and the United States. We find that while some of these frameworks recognize accumulation of privacy harms, they largely emphasize individual control. As a result, current approaches risk privacy harms which may not be as obvious, but can still lead to leakage of sensitive information.

References

Dinur, I., & Nissim, K. (2003). Revealing information while preserving privacy. In *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (pp. 202–210). Association for Computing Machinery. https://doi.org/10.1145/773153.773173

Dwork, C., & Roth, A. (2014). *The algorithmic foundations of differential privacy*. *Foundations and Trends in Theoretical Computer Science, 9*(3–4), 211–407. https://doi.org/10.1561/0400000042

Gordon-Tapiero, A., Ligett, K., & Nissim, K. (2025). *On the rival nature of data: Tech and policy implications*. In *CS and LAW 2025 – Proceedings of the 2025 Symposium on Computer Science and Law* (pp. 17–25). Association for Computing Machinery. https://doi.org/10.1145/3709025.3712211

Varian, H. R. (2018). *Artificial intelligence, economics, and industrial organization* (Working Paper No. 24839). National Bureau of Economic Research. https://www.nber.org/papers/w24839

Veldkamp, L., & Chung, C. (2024). Data and the aggregate economy. *Journal of Economic Literature, 62*(2), 458–484. https://www.aeaweb.org/articles?id=10.1257/jel.20221580