Coding Justice: Architectural Shift from Ex-Post Punishment to Ex-Ante Design in Privacy-Preserving AI

Linyu Zhang

Xi'an Jiaotong University

Abstract

The development of Artificial Intelligence relies on large-scale data, which creates significant privacy and compliance challenges. This article examines how privacy-preserving technologies like federated learning and differential privacy reshape legal relationships and responsibility allocation through their technical architectures. We argue that these technologies do more than protect privacy; they enable a legal paradigm shift from ex post punishment to ex ante design. This shift clearly redefines the responsibilities of data controllers and processors, laying the groundwork for a new, technology-enabled legal governance framework.

Introduction

The rapid development of artificial intelligence, especially large-scale models, has expanded its use in areas like healthcare, finance, and public services. However, a clear tension has emerged between data-driven AI systems and strict data privacy laws such as the EU’ s GDPR and China’ s Personal Information Protection Law (PIPL). This conflict is particularly evident in fields like healthcare and education, which handle highly sensitive information. Here, organizations face the challenge of using AI to improve services while also ensuring data privacy and security. Traditional legal compliance methods rely mainly on contracts and after-the-fact accountability. These approaches often fall short when dealing with the complex and non-transparent data processing practices of modern AI systems. ***In response***, a range of privacy-preserving AI technologies—such as split learning and differential privacy—has emerged. These methods aim to achieve “using data without exposing it” and represent a key research direction today. This paper argues that the value of these technologies goes beyond providing privacy protection tools. More importantly, they reshape legal relationships and responsibility allocation in data processing through architectural redesign. In doing so, they offer a new compliance paradigm for the AI era.

Main contributions

➢ *Articulating a Foundational Paradigm Shift in Legal Governance*

This work posits that privacy-preserving technologies—such as federated learning and split learning—act not merely as tools but as ***architects of a new legal paradigm***. They fundamentally transition privacy compliance from a reactive model reliant on ex-post punishment and contractual liability to a proactive model of ex-ante prevention engineered directly into system architectures.

This signifies a pivotal shift where ***legal compliance is encoded into technical design***, transforming it from an external subject of audit into an inherent, operational property of the system itself.

➢ *Developing a Systematic Framework that Maps Technical Boundaries to Legal Principles*

We provide a coherent analytical framework that systematically demonstrates how the core technical construct of ***"technical boundaries"*** concretely implements fundamental data protection principles. This framework bridges the technical-legal divide by mapping specific architectural features to abstract legal mandates:

- ***Data Minimization*** is achieved by design through local training (FL) or the transmission of only intermediate features (Split Learning).
- ***Privacy by Design*** is embodied by the provable, mathematical guarantees of Differential Privacy.
- ***Purpose Limitation and Security Obligations*** are enforced by technologies like Homomorphic Encryption, which enable computation while physically preventing access to raw data.

This contribution demonstrates that abstract legal principles can be translated into executable, verifiable engineering specifications.

➢ *Proposing a Novel, Actionable Scheme for Reallocating Liability Based on Technical Architecture*

Our most practical contribution is a novel liability framework derived from the aforementioned "technical boundaries." It offers an objective standard to resolve the persistent challenge of "ambiguous liability in data circulation":

- The ***Data Controller's*** liability is anchored at the source, focusing on ensuring local data security and fulfilling transparency obligations.
- The ***Algorithm Developer's/Processor's*** liability is strictly confined to the domain within the technical boundary, limited to processing the received encrypted or de-identified information. Any attempt to cross this boundary to reconstruct raw data constitutes a per se violation.

This architecture-based scheme significantly enhances ***the practicality of enforcement and adjudication***, providing market participants with stable and predictable compliance expectations.

Conclusion

In summary, our research elucidates the profound legal significance of privacy-preserving AI: it is re-engineering the logic of compliance. We articulate the overarching paradigm shift from ex-post to ex-ante governance, develop a structured framework to understand how technology encodes the law, and derive a consequent, clearer, and more equitable model for allocating liability. We argue this provides a critical pathway for realizing data utility within a robust and innovation-friendly regulatory framework. The future improvement of AI governance relies on a deep integration of technical insight and legal wisdom. Constructing a legal environment that fosters innovation while protecting fundamental rights and clarifying responsibilities will be a key focus for future work.