

ARTICLE

Statutory Obsolescence in the Age of Innovation: A Few Thoughts about GDPR

Florence G'sell

The Network Law Review is pleased to present a special issue entitled “The Law & Technology & Economics of AI.” This issue brings together multiple disciplines around a central question: What kind of governance does AI demand? A workshop with all the contributors took place on May 22–23, 2025, in Hong Kong, hosted by Adrian Kuenzler (HKU Law School), Thibault Schrepel (Vrije Universiteit Amsterdam), and Volker Stocker (Weizenbaum Institute). They also serve as the editors.

Abstract. This article examines the problem of statutory obsolescence in the regulation of rapidly evolving technologies, with a focus on GDPR and generative AI. It shows how core GDPR provisions on lawful processing, accuracy, and erasure prove difficult—if not impossible—to apply to AI systems, generating legal uncertainty and divergent national enforcement. The analysis highlights how comprehensive, principle-based instruments can quickly become inadequate in fast-moving technological domains. Drawing lessons from the GDPR, the article reflects on the need for more adaptive, flexible, and responsive regulatory approaches in the technological age.

*

Throughout legal history, jurists have grappled with the challenge of adapting established legal frameworks to evolving societal conditions. Traditionally, these challenges were regarded as an inevitable imperfection of the legislative process—one

that careful drafting and judicial interpretation could partly mitigate.¹ In France, Napoleon's Civil Code of 1804 gradually became outdated amid industrialization, prompting the legal community to advocate for judicial discretion and interpretive flexibility as mechanisms for preserving coherence and adaptability.²

This problem has acquired new urgency in the digital age. Technological developments now unfold at such speed that statutes may become impracticable within only a few years of adoption. Legal principles can quickly become misaligned with technological realities, while compliance obligations may soon prove technically infeasible. How can legal norms regulate technologies whose characteristics evolve almost weekly? How can legislators anticipate and mitigate the risks posed by such rapidly shifting functionalities and applications? Can courts adapt effectively when judges may lack the technical expertise required to grasp complex technological issues?

This article examines a recurrent critique of comprehensive regulatory interventions in the technological domain: that legislators, unable to foresee the trajectory of technological development, enact legal norms that risk becoming obsolete or ill-suited almost immediately. The application of the General Data Protection Regulation (GDPR)³ to generative AI provides a clear illustration. Adopted in 2016, before the emergence of foundation models and generative AI, the GDPR already shows signs of strain. Less than a decade later, applying key provisions to AI raises significant difficulties. This tension highlights the structural difficulty of reconciling broad, principle-based legal instruments with technologies that evolve at extraordinary speed.

The article proceeds as follows: the first section analyses the principal challenges that generative AI poses for data protection; the second examines GDPR provisions that are especially difficult to apply in this context; the third reflects on the broader implications of these difficulties and draws lessons for regulatory design.

¹ In the United States, Guido Calabresi discussed the problem of "statutory obsolescence," whereby laws gradually become outdated, irrelevant, or ineffective. He argued that courts should have the authority to revise—or, in exceptional cases, to sunset—statutory provisions that no longer correspond to prevailing social and legal realities. Guido Calabresi, *A Common Law for the Age of Statutes* (1982), Harvard University Press.

² Christophe Jamin and Philippe Jestaz, *La doctrine* (2014), PUF, collection *Méthodes du droit*.

³ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4.5.2016, pp. 1–88.

1. Data protection challenges posed by generative AI

Personal data encompasses any information that directly or indirectly identifies a living individual. Such data may appear at multiple stages of the AI lifecycle: in training datasets, user prompts, and model outputs. An additional question, still unresolved both legally and technically, concerns whether the trained model itself may be regarded as incorporating personal data.

The processing of personal data in these contexts gives rise to substantial legal challenges. A central concern is that personal data may be collected and used without the knowledge or consent of the individuals concerned. Developers generally disclose little information about the provenance or composition of training corpora, which are often assembled through large-scale web scraping and may include personal information.⁴ Advanced AI models can also “memorize” specific elements of their training data, which may be regurgitated in generated outputs.⁵ Even without direct memorization, personal data may be inferred through patterns and associations encoded in the model parameters. In a model inversion attack, for instance, an adversary may employ a secondary model to reconstruct or infer training data from a target model’s outputs.⁶

At the output level, generative AI systems may generate factually inaccurate or misleading content with tangible consequences for individuals.⁷ A notable case involved a law professor who discovered that ChatGPT had fabricated a sexual harassment case and erroneously identified him as one of the accused.⁸ Incidents of this kind underscore the potential for reputational harm.

⁴ Florence G’sell, *Regulating under Uncertainty: Governance Options for Generative AI* (2024), Stanford Cyber Policy Center, 40-41 and 97-99.

⁵ Milad Nasr et al., « Scalable Extraction of Training Data from (Production) Language Models », arXiv (Nov. 28, 2023), <https://arxiv.org/pdf/2311.17035>

⁶ Nicholas Carlini et al., « Extracting Training Data from Large Language Models », arXiv (June 15, 2021), <https://arxiv.org/pdf/2012.07805>; Another type of attack is membership inference, which leverages a model’s behavior to infer whether particular data points were part of its training dataset, Nicholas Carlini et al., “Membership Inference Attacks From First Principles,” 2022 IEEE Symposium on Security and Privacy (2022), pp. 1897-1914.

⁷ OpenAI has disclosed concerning data regarding the frequency of errors generated by certain of its models: OpenAI, Open AI o3 and o4-mini System Card (April 16, 2025) <https://cdn.openai.com/pdf/2221c875-02dc-4789-800b-e7758f3722c1/o3-and-o4-mini-system-card.pdf>

⁸ Pranshu Verma & Will Oremus, ChatGPT invented a sexual harassment scandal and named a real law prof as the accused, Wash. Post (Apr. 5, 2023), <https://www.washingtonpost.com/technology/2023/04/05/chatgpt-lies/>

2. GDPR compliance challenges for AI providers

Although the GDPR was adopted relatively recently, its provisions present significant challenges when applied to generative AI, generating substantial legal uncertainty. Three areas illustrate these tensions particularly clearly: the lawfulness of processing, the principle of accuracy, and the right to erasure. Each highlights the broader difficulty of applying principle-based legislation to rapidly evolving technologies.

2.1. Lawfulness of processing

Under Article 6 GDPR, the processing of personal data must rest on a valid legal basis throughout the lifecycle of an AI model, from pre-training to fine-tuning. The two Article 6 bases that most plausibly apply to AI model development are: (i) the explicit consent of the data subject, or (ii) the necessity of processing for the legitimate interests of the controller. Both prove problematic in the AI context.

Obtaining explicit consent is virtually impossible when training data are scraped from the web and concern millions of individuals. GDPR further requires that consent be specific, informed, and tied to each distinct purpose of processing – standards incompatible with the evolving purposes of general-purpose models and the unpredictability of downstream applications. In practice, consent can realistically be secured only at the point of registration for a specific service, for example when individuals agree that the data they provide during interactions may be processed.

Reliance on legitimate interest under Article 6(1)(f) also presents challenges. Processing is permitted when necessary to pursue the data controller's legitimate interests, provided these interests are not overridden by the rights of the data subject. In 2023, the CJEU ruled that Meta could not invoke legitimate interest to justify tracking and profiling for behavioural advertising, despite offering its service free of charge.⁹ This judgment raised doubts about the applicability of legitimate interest to AI development.¹⁰

⁹ Case C-252/21, *Bundeskartellamt v Meta Platforms Ireland Ltd*, ECLI:EU:C:2023:537 (CJEU, 4 JULY 2023)

¹⁰ For instance, in its decision dated 2 November 2024, which was accompanied by a press release published on 20 December 2024, the Garante found that OpenAI had failed to establish a valid legal basis for the processing of personal data at the time of the model's public release and during the pre-launch training of ChatGPT. Garante per la protezione dei dati personali, *Decision about OpenAI*, Press release (Dec 20, 2024)

In December 2024, the European Data Protection Board (EDPB) issued an Opinion¹¹ recognizing that AI development may rely on legitimate interest, provided developers conduct a rigorous three-part assessment: (i) identify a legitimate aim, such as creating a conversational agent to assist users; (ii) confirm that processing is necessary to achieve that objective and that less or no personal data cannot reasonably achieve the goal; and (iii) ensure the legitimate interest pursued is not overridden by the rights of data subjects. Shortly thereafter, Meta announced that it would begin training AI models on publicly shared data from adults across its EU platforms,¹² reversing its June 2024 suspension following objections from the Irish Data Protection Commission (DPC).¹³ The announcement drew immediate criticism from privacy advocates,¹⁴ highlighting the persistent legal and regulatory uncertainties regarding the use of legitimate interest as a lawful basis for AI development.

2.2. Accuracy

Article 5 requires that personal data be “accurate and, where necessary, kept up to date,” and that “every reasonable step” be taken to ensure the prompt rectification or erasure of inaccurate data. Article 16, titled “Right to rectification,” grants individuals the right to have inaccurate personal data corrected and incomplete data completed. These provisions pose substantial challenges for generative AI.

At the training stage, the scale and heterogeneity of datasets—particularly those sourced through web scraping—render it impossible to verify accuracy or systematically correct errors. At the output stage, ensuring accuracy is even more difficult, as generative AI produces probabilistic outputs that may include fabricated or erroneous content. Such outputs cannot be easily anticipated, controlled, or corrected by either developers or end users.

Recent complaints illustrate these difficulties. One complaint against OpenAI claimed that ChatGPT repeatedly returned incorrect information about the

¹¹ EDPB, Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models (Dec 17, 2024), https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf

¹² Reuters, “Meta to use public posts, AI interactions to train models in EU” (April 14, 2025), <https://www.reuters.com/technology/artificial-intelligence/meta-use-public-posts-ai-interactions-train-models-eu-2025-04-14/>

¹³ Reuters, “Meta pauses AI models launch in Europe due to Irish request” (June 14, 2024), <https://www.reuters.com/technology/artificial-intelligence/meta-will-not-launch-meta-ai-europe-now-2024-06-14/>

¹⁴ NOYB, “noyb sends Meta ‘cease and desist’ letter over AI training. European Class Action as potential next step” (May 14, 2025), <https://noyb.eu/en/noyb-sends-meta-cease-and-desist-letter-over-ai-training-european-class-action-potential-next-step>

complainant's date of birth, arguing that "as long as ChatGPT keeps showing inaccurate data," OpenAI fails to comply with Article 5(1)(d).¹⁵ OpenAI responded that the only effective mitigation would be to prevent the generation of any personal data relating to the complainant, but argued that, since he is a public figure, doing so would unjustifiably restrict freedom of expression and access to information. In another case before the Norwegian Data Protection Authority, ChatGPT allegedly generated content falsely accusing an individual of murdering two of his children and attempting to kill the third. The privacy rights organization, NOYB, supporting the complaint, argued that developers are obliged to take "every reasonable" step to ensure that inaccurate personal data are "erased or rectified without delay."¹⁶ Yet because it is neither technically possible to prevent generative AI from producing inaccuracies, nor feasible to systematically correct outputs *ex post*, compliance appears unattainable.

The EDPB acknowledged this tension in its May 2024 report, emphasizing that "a difference should be made between input and output data."¹⁷ It further noted that "the purpose of the data processing is to train ChatGPT and not necessarily to provide factually accurate information." However, the EDPB stressed that developers and providers must transparently communicate the mechanisms behind output generation and the inherent limitations of reliability. This includes explicitly disclosing that outputs, while syntactically correct, may contain bias or fabricated content—a practice already generally followed by AI providers.

2.3. Erasure

Article 17 of the GDPR grants individuals the right to erasure, allowing them to request the deletion of their personal data without undue delay. This raises a central question: can a request for erasure concern an AI model that has already been trained? Specifically, once personal data has been used in training, is the model a repository of personal data, or does the training process effectively render the data anonymous?

In July 2024, the data protection authority of the German state of Hamburg published a paper arguing that general-purpose AI models, once trained, do not

¹⁵ NOYB's Complaint against OpenAI, Case C-078 (April 29, 2024), https://noyb.eu/sites/default/files/2024-04/OpenAI%20Complaint_EN_redacted.pdf

¹⁶ NOYB's Complaint against OpenAI, case C082 (March 20, 2025), https://noyb.eu/sites/default/files/2025-03/OpenAI_complaint_redacted.pdf.

¹⁷ EDPB, Report on the work undertaken by the ChatGPT Taskforce (May 23, 2024), https://www.edpb.europa.eu/system/files/2024-05/edpb_20240523_report_chatgpt_taskforce_en.pdf

themselves constitute the processing of personal data.¹⁸ The paper explained that the tokens and numerical values comprising a model’s internal parameters do not meaningfully correspond to information about identifiable individuals. Under this interpretation, the model could be considered anonymous, although subsequent interactions may still involve personal data if user inputs or model outputs contain such information.

By contrast, the EDPB’s December 2024 Opinion asserted that AI models can, in fact, contain personal data, and whether they do must be assessed on a case-by-case basis. A model trained with personal data may be considered anonymous only if there is a negligible risk that: (i) personal data used during training could be extracted, including through probabilistic methods, or (ii) such data could be unintentionally or intentionally disclosed through user queries. In other words, both the risk of deliberate attacks, such as model inversion, and the risk of inadvertent “regurgitation” of personal data in model outputs must be negligible. The EDPB provided a non-exhaustive, non-prescriptive list of factors relevant to evaluating whether an AI model may be considered anonymous.¹⁹

In any case, once an AI model has been trained or fine-tuned on personal data, reversing the effects of that training appears technically unfeasible. In practice, compliance with a data erasure request for a trained model would require discarding the existing model and retraining a new one on a dataset excluding the individual’s data. An emerging alternative, ‘machine unlearning,’²⁰ aims to remove the influence of specific data from a trained model. Both approaches presuppose that the relevant data can be accurately identified and isolated—a task that is particularly challenging in large-scale datasets. It remains uncertain whether these methods could satisfy GDPR requirements.

¹⁸ Hamburg Commissioner for Data Protection and Freedom of Information (HmbBfDI), “Discussion Paper: Large Language Models and Personal Data” (July 15, 2024), https://datenschutz-hamburg.de/fileadmin/user_upload/HmbBfDI/Datenschutz/Informationen/240715_Discussion_Paper_Hamburg_DPA_KI_Models.pdf

¹⁹ These include: limiting the collection of personal data or applying pseudonymization or filtering before training; employing privacy-preserving techniques during model training, such as differential privacy; implementing safeguards to prevent the model from outputting personal data; establishing strong engineering governance and maintaining documentation for audit purposes; and conducting rigorous testing against known attack vectors, such as attribute and membership inference, data exfiltration, training data regurgitation, model inversion, and reconstruction attacks.

²⁰ G’sell, *Regulating under Uncertainty*, fn 5, pp. 160-161 ;A.F. Cooper et al. “Machine Unlearning Doesn’t Do What You Think: Lessons for Generative AI Policy, Research, and Practice.” *ArXiv* (2024), <https://arxiv.org/abs/2412.06966>.

3. Lessons from the GDPR example

The preceding examples demonstrate that the GDPR is ill-suited to the structural characteristics of the most recent AI models and advanced generative AI systems. This regulatory mismatch generates considerable legal uncertainty and, in practice, has led to divergent outcomes across the European Union.

3.1. Legal uncertainty

At present, substantive reform of the GDPR appears unlikely in the near future,²¹ and EDPB guidance, while technically informed, offers limited legal certainty. This uncertainty stems primarily from the gap between the GDPR's conceptual framework and the technological realities and complexity of AI models. The EDPB implicitly acknowledged this uncertainty by noting that AI models may or may not be considered anonymous, with such determinations requiring case-by-case assessment. Concerning the legal grounds for processing, its interpretation of "legitimate interest" as a potential basis demands complex, context-specific evaluations. This ambiguity creates significant potential for litigation, as illustrated by the ongoing claim against Meta regarding its decision to train AI models on user data. On the principle of accuracy, the EDPB's distinction between training data and AI outputs allows for a flexible application of Article 5 GDPR. Nonetheless, because it maintains that the principle of accuracy continues to apply, the extent to which AI outputs may be exempt from strict compliance with Article 5(1)(d) remains unresolved.

Meanwhile, national data protection authorities have adopted their own interpretations and policies, occasionally imposing particularly stringent measures. For instance, since 2023, the Italian Data Protection Authority (Garante) has strictly enforced GDPR compliance with respect to AI models operating in Italy, issuing a temporary restriction on OpenAI's ChatGPT,²² ordering the suspension of Replika due to risks to minors and unlawful data processing,²³ and blocking DeepSeek on the

²¹ The sole modification directly affecting the GDPR in the European Commission's "Fourth Omnibus" reform package, launched on 21 May 2025, would exempt companies with fewer than 750 employees from the obligation to maintain a record of processing activities; In a separate development, on 16 June 2025 the European Commission and the European Parliament announced agreement on reforms to the GDPR's cross-border enforcement procedure in order to streamline and harmonise enforcement in cases involving multiple member states.

²² Garante, "Artificial intelligence: stop to ChatGPT by the Italian SA, Personal data is collected unlawfully, no age verification system is in place for children", (March 31, 2023), <https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9870847#english>;

²³ Garante, Decision of April 10, 2025, <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10130115#english>

grounds of inadequate transparency in data collection and storage practices.²⁴ Such rigorous measures, however, present challenges: they are limited to a single EU Member State and can restrict access to advanced technologies for both individuals and businesses.

3.2. Towards new regulatory approaches?

These developments raise a broader question: should existing regulatory approaches be redesigned to keep pace with the rapid and continuous evolution of emerging technologies? In an era of swift technological change, reliance on comprehensive legislative instruments that enshrine broad principles and legal concepts risks producing rules that quickly become obsolete or impracticable. Less than a decade after the adoption of the GDPR, several of its key provisions appear misaligned with the realities of AI, and the competent authorities face significant challenges in determining a workable interpretation of the principles it enshrines. In such a fast-moving environment, one-size-fits-all regulations may need to yield to more concrete, adaptable, and technologically responsive standards.

With the AI Act, the European legislator appears to have pursued a more pragmatic approach. The EU AI Office, established within the European Commission, oversees enforcement and implementation across Member States and is supported by advisory bodies designed to facilitate a flexible and responsive application of its principles.²⁵ The Scientific Panel, composed of independent AI experts, provides technical guidance, particularly on systemic risks posed by general-purpose AI models. The Advisory Forum, bringing together diverse stakeholders—including industry, SMEs, startups, academia, and civil society—offers broader input, ensuring a more inclusive and adaptive regulatory process. These mechanisms aim to align standards closely with technological realities; nevertheless, their ultimate effectiveness remains uncertain.

Overall, it is doubtful that such adaptive mechanisms alone can resolve the problem of statutory obsolescence, which requires a fundamental rethinking of lawmaking in the technological age. Should legislators continue to rely on broad regulatory instruments rooted in legal concepts whose interpretation becomes increasingly difficult amid rapid technological change? What role should courts play in this evolving context? Might soft law and self-regulation offer a better approach for particularly complex technologies? While this article does not provide definitive answers, the GDPR example underscores the urgency of rethinking regulatory

²⁴ Garante, Decision of January 30, 2025, <https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/10098477#english>

²⁵ G'sell, *Regulating under Uncertainty*, fn 5, pp. 229-230.

design to maintain coherence and effectiveness in a fast-evolving technological environment.

Florence G'sell