Harvard Data Science Review • Special Issue 5: Grappling With the Generative AI Revolution

Data Protection and Generative AI: Inconclusive Answers to Unresolved Questions

Romina Garrido[1,2]

[1]**GobLab, School of Government, Adolfo Ibáñez University Santiago, Chile,**
[2]**Prieto Abogados, Vitacura, Chile**

The MIT Press

Published on: May 31, 2024

DOI: https://doi.org/10.1162/99608f92.91162b2e

ABSTRACT

The article briefly discusses the challenges of data protection and the possible responses that this discipline can provide to generative AI regulations.

Keywords: generative AI, data protection, algorithms, privacy

---

Generative AI refers to those AI tools capable of creating text, images, music, audio, and videos. These systems use various machine learning techniques that draw inferences from patterns and relationships contained in the large data sets on which they are trained. Such data are primarily produced by people or emanate from their digitally captured behavior (Google Cloud, 2023). Many of these tools are multipurpose with capacity to perform several different tasks depending on the human prompt entered (e.g., making lists, providing summaries, answering questions, producing classifications). The AI chatbot, ChatGPT, is an example of this. It is a large language model that erupted onto the scene seemingly out of nowhere,[1] throwing discussions of AI regulation and governance into almost immediate upheaval.

Much of the turmoil arose from the fact that generative AI systems like ChatGPT do not have a single defined purpose and hence present an expansive risk surface that is difficult to manage and govern. Clear limitations on purpose would help existing regulations, and the regulators who enforce them, to better frame how to oversee these technologies and how to handle their risks. A good example of this governance quandary arises in the difficulty of applying personal data protection regulations to the production and use of these systems. Insofar as they process personal data, generative AI systems are subject to data protection laws in jurisdictions where these laws are in force, and yet there is, up to now, no conclusive answer as to how these regulations should be applied. In confronting whether such systems conform to the demands of data protection law, we are left with only inconclusive answers to unresolved questions.

The right to data protection was born precisely as a response to the impact of data processing on fundamental rights, more than 40 years ago. Since their inception, data protection regimes have gained strength as regulatory tools that confront the myriad ways the use of digital technologies can offend individuals' right to control their personal data. Personal data refers to any information that is capable of identifying a person. Data protection rights include the right to know the practices and policies of data processors, who handle personal data. Where AI systems process personal data in a fully automated way and have a "legal or similarly significant effect" on a person, affected individuals have a right to transparency and explanation, including a right to meaningful information about the logic behind the system's outputs. Data protection establishes rules to carry out the processing of personal data in lawful, transparent, accountable, and fair ways, setting guidelines for those who process any kind of personal data and making them responsible for any damage that unlawful actions may cause or for not complying with the established standards.

The European data protection framework, the General Data Protection Regulation of the European Union (GDPR), contains these types of provisions, which have also been permeating Latin American laws, for instance, in Uruguay, Brazil, Panama, Argentina, and in a draft bill in Chile.

The GDPR, and laws that align with it, have at least partly provided a response to AI-based decision-making. This response has included the codification of different transparency and accountability obligations such as the requirement to notify data subjects of the existence of automated decisions; the requirement to provide access to information about how the data processing led to a specific decision, in a concise, transparent, understandable manner and in clear language (i.e., an explanation of the logic behind the system's output); information regarding those responsible; and the right of data subjects to challenge solely automated decisions.[2]

The interpretation and application of the mechanisms described in the previous paragraph have generated quite an interesting legal debate regarding the existence (or not) of a new right to explanation. For some, the GDPR contains a significant set of rules on algorithmic transparency, bringing clarity to the governance of the design and use of computer-based algorithms (Kaminski, 2021), while others have been more skeptical, stating that regulations have several notable limitations that prevent us from talking about a right of explanation at all (Wachter et al., 2017).

Aspects of transparency and explanation are, however, only part of what is relevant when trying to understand how generative AI can affect people and their data protection rights. Data protection regulations oblige data controllers both to provide a justification of the legal basis for the processing of personal data and to provide mechanisms to enable interested parties, including nonusers, to exercise their data protections rights (e.g., the right to the rectification of inaccurate data). The Italian data protection authority acted from these premises, when, in April 2023, it required OpenAI to comply with disclosure requirements and to provide justification for the legal basis of the processing of the personal data of affected parties for training ChatGPT (Garante, 2023).

This example suggests that the debate could go much deeper, if, for example, one seeks to ask and answer further data protection questions about lawfulness and purpose: Is it legal to use personal data, such as human-generated text and images, that have been publicly shared on the Internet in one specific context for a different one (e.g., using scraped social media posts to train generative AI models)?

In 1997 Helen Nissenbaum developed a theoretical framework for the protection of private information in advertising scenarios, which she called "Privacy as Contextual Integrity" (Nissenbaum, 2004). This theory demands that information should not be used or shared outside its original context. Nissenbaum's framework sought to provide an alternative, more communicative centered way to understand legitimate data use that emphasized the importance of the people's reasonable expectations to contextually bound privacy in public. From the contextual integrity perspective, privacy is a matter of limits. It is about the extent to which it is

permissible to deploy or extend the use of certain personal data outside the bounds of its originating context; when these limits are violated, invasions of privacy occur (Sánchez, 2016).

In the language of data protection laws, this notion of privacy is partially assimilated into one of the concepts that allow the weighing of legitimate interest as a basis of legality other than consent: the 'compatible purpose.' This is one of the aspects to consider to enable the secondary use of data. It requires considering the context of the original processing and the subsequent purposes of use, which in general will be different from the original purposes. From a more stringent perspective, the use of personal data by generative AI violates contextual integrity (Gonzalo, 2023)—as many cases of data use in social networks show (Noain Sánchez, 2015)—since it does not respect the relevance or expectation of use of the original information being shared, nor the original purpose or flow of information.

We cannot, however, deny that many times our personal data must be accessible to and processed by others as part of the cost (and value) of acting in society. Privacy in public—and today on the web, as a digital public square—is based on complex and variable conditions that are dependent on undefined scenarios across space, time, and use. More creative solutions will be necessary to make the benefits of new technologies compatible with personal rights. These solutions are beginning their journey hand in hand with the challenges of generative AI and so-called 'frontier models.' Progress is already being made in addressing some of these challenges. For example, regulatory and governance proposals have been advanced that point toward permanent risk management, transparency, and quality control mechanisms and measures to enforce our rights linked to personal data. Even so, difficulties surrounding the use of personal data in the training and operation of generative AI technologies demand a more flexible, innovation-enabling approach because, at present, there are available only broadly inconclusive answers to unresolved questions.

Disclosure Statement

Romina Garrido has no financial or non-financial disclosures to share for this article.

References

Duran, I. (2023, December 7). Sam Altman es nombrado CEO del año por la revista TIME [Sam Altman named CEO of the year by TIME Magazine]. *Infobae*. https://www.infobae.com/tecno/2023/12/07/sam-altman-es-nombrado-ceo-del-ano-por-la-revista-time/

Garante. (2023). *ChatGPT: Garante privacy, limitazione provvisoria sospesa se OpenAI adotterà le misure richieste. L'Autorità ha dato tempo alla società fino al 30 aprile per mettersi in regola* [Privacy guarantor, temporary limitation suspended if OpenAI adopts the requested measures. The Authority has given the company until April 30 to comply]. Garante. https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9874751

Gonzalo, M. (2023). *Las lagunas en torno a ChatGPT y la privacidad: Vacío legal o pesadilla de protección de datos* [The loopholes around ChatGPT and privacy: Legal loophole or data protection nightmare]. Newtral. https://www.newtral.es/chatgpt-privacidad-ia-generativas-rgpd/20230326/

Google Cloud. (2023). *Generative AI use cases*. Google. https://cloud.google.com/use-cases/generative-ai?hl

Kaminski, M. E. (2021). The right to explanation, explained. In *Research handbook on information law and governance* (pp. 278–299). Edward Elgar.

Malik, A. (2023, November 6). OpenAI's ChatGPT now has 100 million weekly active users. *TechCrunch*. https://techcrunch.com/2023/11/06/openais-chatgpt-now-has-100-million-weekly-active-users/

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, *79*, 119.

Noain Sánchez, A. (2015). *La privacidad como integridad contextual y su aplicación a las redes sociales* [Privacy as contextual integrity and its application to social media]. *ZER: Revista De Estudios De Comunicación = Komunikazio Ikasketen Aldizkaria*, *20*(39). https://doi.org/10.1387/zer.15531

Sánchez, A. N. (2016). *La protección de la intimidad y vida privada en Internet: Los flujos de información y la integridad contextual en las redes sociales (2004-2014)* [The protection of privacy and private life on the Internet: Information flows and contextual integrity in social networks (2004–2014)] [Unpublished doctoral dissertation]. Universidad Complutense de Madrid.

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law*, *7*(2), 76–99.

Footnotes

1. ChatGPT reaches more than 10 million daily users in less than 40 days after its launch in November 2022 (Duran, 2023). OpenAI's ChatGPT now has 100 million weekly active users (Malik, 2023). ↩

2. Algorithmic transparency is a complex concept that, in fact, has diverse criteria. Many different parameters and transparency requirements can be developed, but ultimately their effective execution must be determined on a case-by-case basis. This can be seen in a study of algorithmic transparency in the Chilean public sector. https://goblab.uai.cl/transparencia-algoritmica-en-el-sector-publico/ ↩