‹ Technology Radar ›

Generative Artificial Intelligence

ANPD

Brazilian Data Protection Authority

‹Technology Radar – short version in English›

n° 1

Generative Artificial Intelligence

Fabiana Faraco Cebrian
Albert França Josuá Costa
Fernanda Nunes Feitosa Barros
Thiago Guimarães Moraes

ANPD
Brasília, DF
Brazil
2024

ANPD
Brazilian Data Protection Authority

Director-President
Waldemar Gonçalves Ortunho Junior

Directors
Arthur Pereira Sabbat
Miriam Wimmer

Authors
General Coordination of Technology and Research
Fabiana Faraco Cebrian
Albert França Josuá Costa
Fernanda Nunes Feitosa Barros
Thiago Guimarães Moraes

1ª edition
Digital publication – PDF
Technology Radar – Short version in English, Number 01, December 2024

ANPD
SCN, Qd. 6, Conj. A,
Ed. Venâncio 3000, Bl. A, 9º andar
Brasília, DF · Brasil · 70716-900
t. +55 61 2025-8101
www.gov.br/anpd

‹ Generative Artificial Intelligence[1] ›

The Brazilian Data Protection Law, *Lei Geral de Proteção de Dados Pessoais (LGPD)*, aims to protect fundamental rights of freedom and privacy and the free development of personality, while being based on economic, technological development and innovation. Thus, technological innovation must be in harmony with the protection of personal data. For adequate protection of data subject rights, it is necessary to understand how data, particularly personal data, is processed in Artificial Intelligence based on generative models.

The generative approach is distinct from other artificial intelligence approaches as it possesses the ability to generate content (data), in contrast to the discriminative approach, which, in summary, allows the system to learn how to make decisions according to the data used. It is important to highlight that the expression Artificial Intelligence is currently used to refer to Artificial Intelligence systems based on Machine Learning. However, Artificial Intelligence is a much broader field than Machine Learning. For the purposes of this text, the terms will be used interchangeably.

For better understanding of this text, the definition of some concepts is necessary. The first concept is **Artificial Intelligence**, which can be defined as the area of human knowledge that studies the development of artificial intelligence systems. An **Artificial Intelligence System** is defined as a machine-based system that, for explicit or implicit objectives, from the input it receives, infers how to generate outputs such as predictions, content, recommendations, or decisions. **Machine Learning** is defined as the ability of a machine to improve its performance in carrying out a task through experience, whereas **Deep Learning** is a variant of Machine Learning that uses multiple layers to solve more complex problems with high accuracy and transforming data in each layer. A **Generative Model** is a probabilistic model capable of generating new data through sampling, while **Prompt** is defined as the text input used to provide instructions to the artificial intelligence model.

1 *This document presents a synthesized version of the publication* ***Technology Radar – Generative Artificial Intelligence (Portuguese)****.*

https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/documentos-de-publicacoes/radar_tecnologico_ia_generativa_anpd.pdf

Finally, the **Artificial Intelligence System Lifecycle** is defined as a framework that describes the evolution and stages of an artificial intelligence system, from the beginning of its development to its deactivation.

Generally, data used to develop an AI system is divided into training and testing sets. The former is used for training the system. In turn, the latter is used to measure its performance. Thus, this approach is called Training and Testing. Furthermore, generative systems present fundamental characteristics: (i) the need for large volumes of data for their training; (ii) the capability of inference that allows the generation of new data similar to training data; and (iii) the adoption of a diverse set of computational techniques, such as transformer architectures and generative adversarial networks.

Such characteristics directly affect the processing of personal data and the principles established in the LGPD. The need for large volumes of data can lead to the processing of both personal and non-personal data. The coexistence of these two types of data increases the risks related to the protection of personal data since it increases the likelihood that personal data will be processed without proper safeguards and that the **principle of necessity** will not be met. Furthermore, the ability to generate new data, or synthetic content, puts the protection of personal data at risk, since the generated synthetic content may prove to be indistinguishable from personal data, by relating to an identified or identifiable natural person, as well as being erroneously associated with these individuals.

Thus, data subjects, data controllers and data processors are faced with AI systems that present significant challenges to the protection of personal data. The risks range from the potential processing of personal data and the generation of synthetic content to the difficulty of ensuring principles such as transparency. These conditions require attention to the principles and rules established by the LGPD and pose challenges related to the risk of violations of fundamental rights.

The relationship between the development of generative artificial intelligence systems and data processing can be evidenced, in a non-

exhaustive way, in four points delimited in this study: (i) **collection**, (ii) **processing,** (iii) **sharing**, and (iv) **elimination**.

One of the ways to **collect** data for the development of generative artificial intelligence systems happens through data scraping from the internet (data scraping or web scraping). Scraping uses dedicated software to navigate the internet and collect data automatically, which may include names, surnames, addresses, email addresses, videos, audios, images, comments, opinions, preferences, among other data and identifiers. The dynamism and speed of new data availability contribute to the scraping and aggregation of data. This, combined with the absence of adequate pre-processing steps for the elimination or anonymization of personal data, increases the risk of improper personal data processing operations. It is important to underscore that the content of public or publicly accessible websites is subject to the rules established in the LGPD.

Processing activities occur throughout the life cycle of the artificial intelligence system, and begin even before the generation of the model, during the creation of the databases that will be used to generate it. During training, the model parameters have their values adjusted to represent the patterns and relationships learned from the training data. In this way, the data is not stored directly in the models, but rather their relationships, making it difficult to identify personal data in the models. However, Generative AI systems allow user interactions in natural language with the trained model to generate responses. Therefore, depending on the form of interaction, instructions and the context informed by the user through the prompt, the responses provided by these systems may contain personal data.

The content generated by the model, although synthetic, may result in the production of inaccurate or false information about an individual. This possibility presents risks to personal data protection, **when confronting principles such as data quality**, as well as the free development of one individual’s personality, including the data subject's image rights.

Data **sharing** is the broadest element and can be considered from different perspectives, such as: (i) data sharing by the system user, who may or may not be the data subject; (ii) the sharing of the outputs of the models; and (iii) the sharing of the pre-trained model with personal data. In the first perspective, the instructions provided by users to the systems may include a variety of information that may present personal data of the user **themselves** or third parties. The data processing agent and the data subject, in many cases, may not be aware of the risks involved in this information sharing. In the second perspective, generative artificial intelligence systems may result on personal data to be shared with third parties, when such data is present in the model outputs. In this case, there is the risk that the shared data will be reused for further purposes, thus requiring to establish of a chain of responsibility among the different agents involved to ensure compliance with the LGPD, even though there are challenges related to this process. The last perspective is the sharing of the artificial intelligence models themselves and, consequently, the built data representation, which may lead to the use of personal data for purposes different from those initially specified.

Finally, the issue of the erasure of personal data in the context of generative AI is analyzed. The elimination of personal data in artificial intelligence systems must consider at least three aspects: (i) the generation of synthetic content; (ii) interaction with the prompt; and (iii) the continuous refinement of the model. This new technological context may result in the possibility of continuous personal data processing and require for new approaches. Thus, there is a challenge in determining what is the period for data retention, as well as whether the principles of purpose specification and necessity have been respected, in addition to difficulties related to the effective revocation of the data subject's consent in generative artificial intelligence systems.

In this brief analysis, it is possible to observe that some principles of the LGPD have relevance and challenges in the context of personal data processing by generative AI systems. Firstly, the **principle of transparency** establishes that information must be clear, precise, and easily accessible to data subjects. However, it is common for data subjects not to be informed about the collection of their data, nor its inclusion in the training sets of the models. In turn, the **principle of**

necessity presents an additional challenge related to the use of large databases in modern generative artificial intelligence systems and meeting the criterion of limiting processing to the minimum necessary to achieve a specified purpose. Even though the principle does not necessarily means a prohibition on using large volumes of data, it involves ensuring that unnecessary personal data is not processed for the purpose defined by the data controller. Finally, the **principle of data quality** prescribes the accuracy, clarity, relevance, and updating of data, implying the adoption of measures to mitigate the risk that AI systems produce inaccurate information about data subjects.

To ensure data protection in the context of generative artificial intelligence one must consider ethical, legal, computational, and sociotechnical perspectives altogether. If not, technological innovation in the area may bring new risks or amplify some that are already known.

‹ References ›

ABNT. Associação Brasileira de Normas Técnicas. ABNT NBR ISO/IEC 22989: 2023. Tecnologia da informação — Inteligência artificial — **Conceitos de inteligência artificial e terminologia**. 1 ed. Rio de Janeiro, 2023.

AEPD. Agencia Española Protección Datos. **Synthetic data and data protection**. 2023. Disponível em https://www.aepd.es/en/prensa-y-comunicacion/blog/synthetic-data-and-data-protection. Acesso em 11 de jan. de 2024.

BRASIL. **Lei Geral de Proteção de Dados**. 2018. Disponível em https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm. Acesso em 11 de jul. de 2023.

CNIL. Commission Nationale de l'Informatique et des Libertés**. AI how-to sheets**. Disponível em: https://www.cnil.fr/en/ai-how-sheets. Acesso em: 02 fev. 2024.

FOSTER, David. **Generative Deep Learning: Teaching Machines to Paint, Write, Compose and Play**, O`Reilly, 2019.

GARTNER. **Gartner Glossary**. Disponível em: https://www.gartner.com/en/information-technology/glossary/foundation-models. Acesso em 02 de fev. de 2024.

GLENSTER, Ann Kristin; GILBERT, Sam. **Policy Brief: Generative AI Report**. University of Cambridge. 42 p. 2023.

MITCHEL, Tom. **Machine Learning**. McGraw-Hill Science, 1997.

OCDE. **OECD Ai Principles overview**. Disponível em: https://oecd.ai/en/ai-principles. Acesso em 17 de abr. de 2024.

OPC. Office of the Privacy Commissioner of Canada**. Joint statement on data scraping and protection of privacy**. 2023. Disponível em: https://www.priv.gc.ca/en/opc-news/speeches/2023/js-dc_20230824/. Acesso em: 02 de fev. 2024.

RANA, Md Shohel; et al. **Deepfake Detection: Systematic Literature Review**. IEEE Access. v. 10, p. 25494 - 25513, fev. 2022.

SOLOVE, Daniel Justin. **Artificial Intelligence and Privacy.** 2024. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4713111. Acesso em 2 de fev. de 2024.

TEFFÉ, Chiara Spadaccini de. **Considerações sobre a proteção do direito à imagem na internet.** Revista de Informação Legislativa: RIL, v. 54, n. 213, p. 173-198, jan./mar. 2017.

www.gov.br/anpd

ANPD