‹ technology radar ›

generative artificial intelligence

ANPD

National Data Protection Agency

‹ technology radar ›

n° 3

generative artificial intelligence

Albert França Josuá Costa
Fabiana Faraco Cebrian
Lucas Costa dos Anjos
Marcelo Santiago Guedes
Thiago Guimarães Moraes

ANPD
Brasília, DF
2026

Cataloging in Publication (CIP)

A2656g

Generative artificial intelligence/ Agência Nacional de Proteção de Dados. – Brasília, DF: ANPD, 2026.

35 p.il.: Color. (Technology Radar; v. 3)

ISBN: 978-65-82658-06-8

Includes bibliographical references.

1. Generative artificial intelligence. 2. Machine learning. 3. Personal data protection. I. Agência Nacional de Proteção de Dados - ANPD. II. Title.

CDU 004.8

Cataloging data prepared by librarian
Angela Halen Claro Franco, CRB1- S023

ANPD – National Data Protection Agency

Director-President
Waldemar Gonçalves Ortunho Junior

Directors
Iagê Zendron Miola
Lorena Giuberti Coutinho
Miriam Wimmer

Project administration
Superintendency of Innovation and Technology (SITEC)
Lucas Costa dos Anjos
Angela Halen Claro Franco
Ewerton Luiz Costadelle

Authors
Albert França Josuá Costa
Fabiana Faraco Cebrian
Lucas Costa dos Anjos
Marcelo Santiago Guedes
Thiago Guimarães Moraes

Translators
Albert França Josuá Costa
Lucas Costa dos Anjos
Gabriela Campos Alkmin
José Pedro Morais de Araujo
Henrique Hermont Barbosa

Graphic design / cover
André Scofano Maia Porto

Desktop publishing
Bruna Assi Hernandes

How to cite this publication:
ANPD. **Generative Artificial Intelligence.** Brasília, DF: ANPD, 2026. (Technology Radar, n. 3). Available at: include document link. Accessed on: include date of access to the document.

ANPD
SCN, Qd. 6, Conj. A, Ed. Venâncio 3000, Bl. A, 9th floor
Brasília, DF · Brazil · Postal Code: 70716-900
Phone: +55 (61) 2017-3338 · www.gov.br/anpd

‹ about the series ›

The "Technology Radar" series is a periodic publication by the ANPD that aims to provide concise overviews of emerging technologies that will impact, or are already impacting, the national and international data protection landscape.

The purpose of the series is to provide relevant information for the debate on data protection in Brazil, with texts structured in an educational manner that is accessible to the wider public, without the intention of exhausting the topics or establishing institutional positions.

For each topic, the main concepts, potentialities, and prospects are addressed, with a special emphasis on data protection within the Brazilian context. ▪

explanatory note

The study on Generative Artificial Intelligence was originally developed by the General Coordination of Technology and Research as an internal resource to support the units of the National Data Protection Agency (ANPD) in 2023. Its initial purpose was to support technical reflections and discussions within the institutional context.

We emphasize that, since its preparation, the technological landscape related to Generative Artificial Intelligence has evolved rapidly. For this reason, this edition of the Technology Radar may contain information that, although relevant at the time of its production, does not fully reflect the most recent developments in the field, such as those discussed in the section on prospects.

We believe that the publication of this edition remains relevant, as it provides a solid foundation and starting point for reflections and challenges related to Generative Artificial Intelligence and the protection of personal data. ▪

‹ summary ›

‹ introduction ›

The creation of intelligent machines has long been a human endeavor, dating back to the 18th century, when the machine called *The Turk* impressed the world by playing chess, supposedly equipped with intelligence. The term *Artificial Intelligence*[1] (AI) was initially coined in the 1950s, during the Dartmouth conference (Dartmouth, c2023). Since then, AI has driven great advances in science and has undergone remarkable internal shifts.

In its early stages, AI was based on the symbolic approach, in which a set of rules is identified by human experts and subsequently translated into logical instructions to be executed by machines. However, this approach presents significant limitations, as the real world is highly complex, making it impossible to manually encode knowledge on a large scale (Russel; Norvig, 2013).

These limitations have led the scientific community to look for mechanisms that would allow machines to, autonomously, learn the rules required to solve the most complex problems, based on historical data from the problems themselves, rather than being explicitly programmed. This search led to the emergence of the field of *Machine Learning*[2] (ML).

As Solove (2024) explains, most of the time the term AI is used today, it is used to refer to *AI systems*[3] based on ML. It is important to bear in mind that AI is a much broader field than ML; however, for the purposes of this study, the expressions will be used interchangeably.

At the beginning of the 21st century, the field of ML underwent exponential growth, leading to the emergence of the subfield known as *Deep Learning*[4] (DL). DL models are widely used in AI applications today. More recently, the emergence and popularization of *generative models*[5] has gained notoriety. These models are part of the DL subfield and have enormous potential for use across various areas, such as the formulation of new pharmaceutical products (Tong *et al.*, 2021).

1 ***Artificial Intelligence*** *is the field of human knowledge that studies the development of artificial intelligence systems.*

2 ***Machine Learning*** *is the ability to improve [a machine's] performance in some task through experience (Mitchell, 1997).*

3 *An* ***AI system*** *is a machine-based system that, for explicit or implicit purposes, infers, from the inputs it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptability after deployment (OECD, c2024).*

4 ***Deep Learning*** *is a variant of Machine Learning that uses multiple layers to solve problems by extracting knowledge from raw data and transforming it in each layer. These layers incrementally obtain high-level features from the raw data, allowing more complex problems to be solved with high accuracy and less dependence on manual adjustments (Gartner, c2024).*

5 ***Generative models*** *describe how a data set is generated in terms of a probabilistic model. This model can generate new data through sampling (Foster, 2019).*

This technical study aims to provide a foundation for the concepts related to generative Artificial Intelligence models, to identify potential risks to privacy and data protection, and to preliminarily relate them to the Brazilian Data Protection Law (LGPD).

‹ key concepts ›

Among the various definitions of Machine Learning (ML), Mitchell (1997) describes it as the ability to improve [a machine's] performance at a task through experience. The emergence of ML has enabled machines to automatically solve more complex problems, such as detecting fraud in credit operations, detecting and classifying tumors in image scans, characterizing elements of hate speech, among other tasks.

Advances in the field of ML have brought to light new terms, such as "language model", "generative model", and "foundation model", which are often used interchangeably[6]. However, although these concepts are related, they have distinct meanings that must be correctly differentiated.

Artificial Neural Network[7] models (Haykin, 1999) are the best known in the ML field, but there are several other models, such as K-Nearest Neighbors (Fix; Hodges, 1951); Decision Trees (Quinlan, 1979) and Support Vector Machines (Cristianini; Shawe-Taylor, 2000).

Regardless of the ML model, its generation generally occurs in two steps: *training*[8] and *testing*[9] (Haykin, 1999).

Generating Machine Learning Models

The primary premise for building machine learning models is the existence of historical data that sufficiently represents the problem to be solved using AI. In simplified terms, the underlying assumption is that what is lacking in knowledge of how to model complex problems is compensated for by an abundance of data.

In a classical approach, these data are divided into training and test sets. The former is used to train the model, while the latter is used to measure its performance. This approach is called Training and Testing (Figure 1).

6 *The distinction between the concept of Language Model and Generative Model will be presented below.*

7 ***Artificial Neural Network*** *is a machine learning model bioinspired in the way the brain processes data (Haykin, 2019).*

8 ***Training:*** *the machine learning model learns the data pattern extracted from a data set.*

9 ***Testing:*** *the model's performance is measured on a different set of data from that used to train the model.*

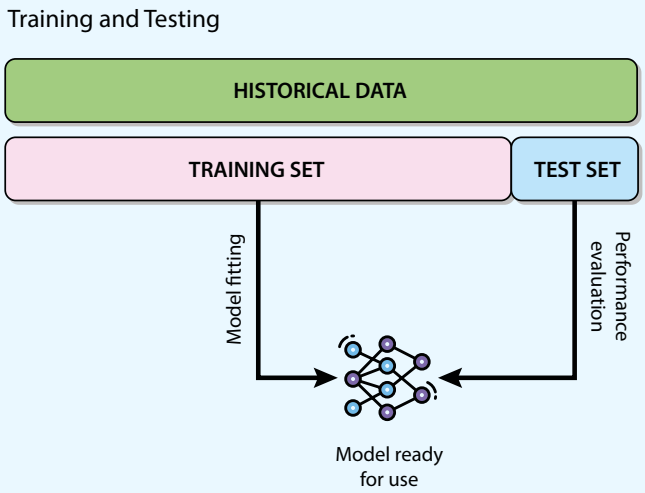

Figure 1 *Model generation using the Training and Testing approach.*

Source: ANPD

Before delving into Generative AI models, it is important to study discriminative models, which are commonly used to classify groups and/or predict trends or behaviors.

Discriminative Model

Generally speaking, ML models are discriminative. This type of model aims to learn, directly from the data, the relationship between the features and what they represent. In short, discriminative models learn a function that predicts a label for each data point.

This decision function is crucial for the tasks performed by discriminative models, such as the classification task, which categorizes data into previously known classes, and the regression task, which predicts the numerical value of a continuous variable based on the data values.

Success in the classification task depends on several factors, ranging from the quality of the data to the careful choice of the algorithm and its parameters. Techniques such as preprocessing and feature selection can be essential for improving model performance, ensuring that it captures relevant information and ignores irrelevant noise.

In turn, mastering the regression problem opens a range of opportunities across various fields. In finance, for example, it enables the prediction of market trends and the assessment of investment risks. In healthcare, it contributes to disease diagnosis, personalized treatment, and the optimization of hospital resource management. In industry, it supports process optimization, demand forecasting, and quality control.

Generative Model

In 2014, Ian Goodfellow published the pioneering paper on *Generative Adversarial Networks* (Goodfellow *et al.*, 2014), which became the starting point for the development of a new type of ML model, termed generative models. Rather than merely learning the decision function of discriminative models, generative models attempt to capture the underlying statistical features of the data to generate new examples that resemble real data.

A generative ML model aims to generate *synthetic data*[10] with the same statistical characteristics as real data. Ideally, the generated synthetic data should be so close to the real data that distinguishing between them would not be possible.

There are currently two main approaches to generative models: the first is based on *Generative Adversarial Networks*[11] (GAN), and the second on *Generative Pre-Trained Transformers* (GPT)[12] (Harrys; Zhu, 2023).

To achieve this objective, models based on Generative Adversarial Networks are composed of two main components: the Generator and the Discriminator. The Generator is responsible for attempting to deceive the Discriminator by generating synthetic data with statistical characteristics similar to those of real data. In turn, the Discriminator seeks to identify which data is real and which is not. The generative model becomes appropriate when the Discriminator can no longer distinguish between real and synthetic data (Figure 2).

10 ***Synthetic data*** *are artificially generated data, in contrast to real data, which come from real-world contexts (AEPD, 2023). The term "synthetic data" is used both in the field of Information Technology (IT) and in the area of privacy and data protection. In the latter, the use of this data is usually associated with privacy-enhancing technologies (PETs). Accordingly, in chapter 3, which discusses the relationship between generative AI and data protection, the term "synthetic content" will be used to refer to the synthetic data generated by generative models. Synthetic content is not intended to anonymize personal data.*

11 ***Generative Adversarial Networks*** *are machine learning models that use the adversarial approach to generate synthetic content.*

12 ***Generative Pre-Trained Transformers*** *are machine learning models that use the encoder-decoder approach to generate synthetic content.*

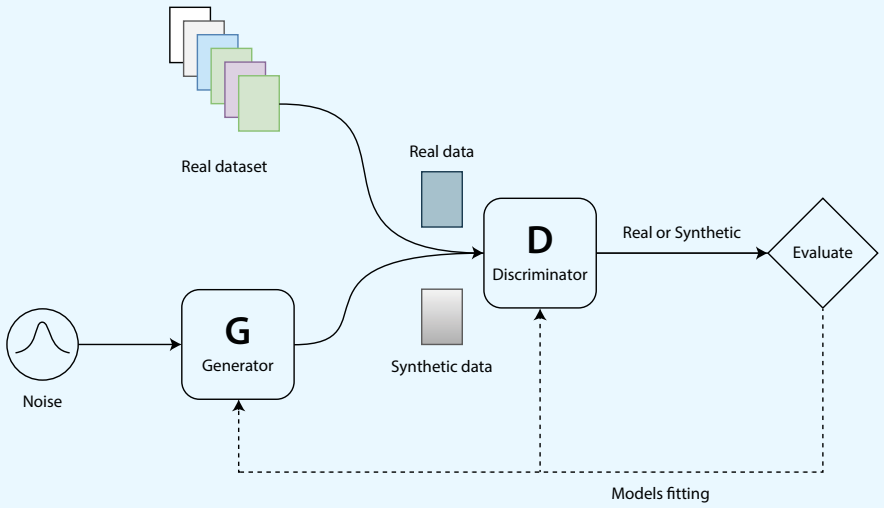

Figure 2 *Structure of a generative model based on generative adversarial networks.*

Source: ANPD

The approach based on Generative Pre-Trained Transformers emerged with the development of two techniques: (i) the transformer architecture (Foster, 2019) and (ii) the attention mechanism (Vaswani *et al.*, 2017).

To achieve the goal of generating data, the transformer architecture is composed of encoder and decoder layers. The encoding layers receive data as input and transform it into a numerical representation known as an *embedding*, which represents the meaning of the input. The decoders, in turn, generate the output based on the information contained in the embedding and on prior context (Figure 3).

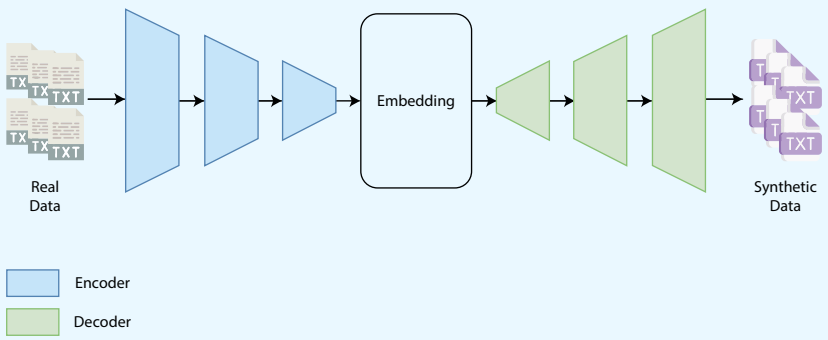

Figure 3 *Structure of a generative model based on pre-trained generative transformers.*

Source: ANPD

For their part, the attention mechanism allows the generative model to focus on specific parts of the real data, rather than treating it uniformly. This selective focus is useful when the relevant elements for the execution of a given task are distributed throughout the input. The mechanism dynamically assigns different weights to the parts of the text based on their relevance to task execution, allowing the more important parts to be prioritized over less important ones (Vaswani *et al.*, 2017).

Two examples are provided to illustrate the concept of the attention mechanism. The first utilizes a computer vision task, as generative models, especially those incorporating attention mechanisms, are also capable of handling various tasks, including computer vision. The second example utilizes a natural language processing task.

Figure 4 illustrates the operation of the attention mechanism within a computer vision task. The model aims to identify which animals are present in the original image (left part of the figure). The lighter regions in the last two images represent the areas to which the model paid greater attention, considering them more important in performing the task. Conversely, the darker regions were considered less important by the attention mechanism.

Figure 4 *Example of an attention mechanism in computer vision.*

Source: Chollet (2021).

Figure 5 illustrates the operation of the attention mechanism within a natural language processing task. The model aims to identify which words merit greater attention as they are deemed more important when fed into a seed word. In the matrix, the rows represent the input words, and the columns represent which words should receive more attention based on each input word observed. For example, when the system observes the word "animal", the model assigns the highest attention scores to "street" (0.17) and "tired" (0.16), whereas "was" receives a lower attention score (0.01).

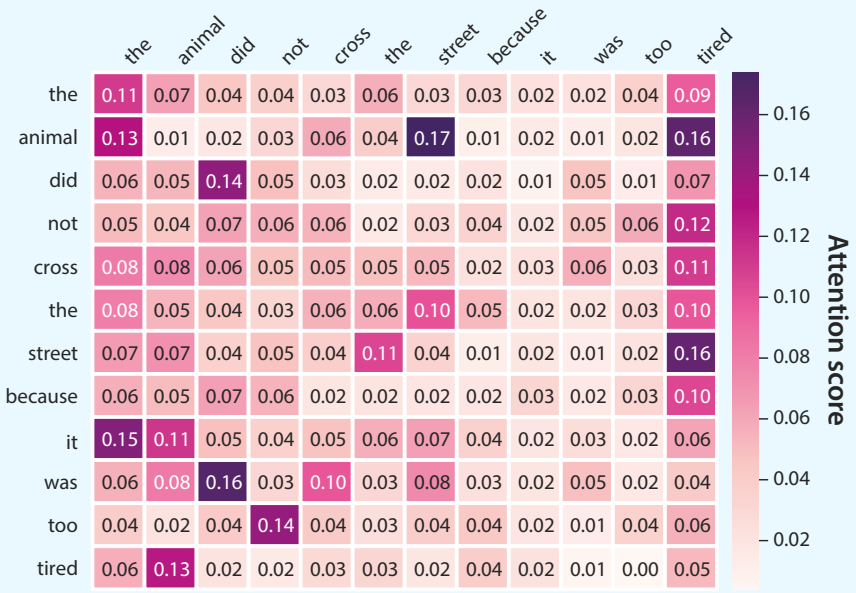

Figure 5 *Example mechanism of attention in natural language processing.*

Source: Adapted from Uszkoreid (2017).

One of the main examples of transformer-based models is known as the *Large Language Model*[13] (LLM). LLMs are trained on massive volumes of textual data, capturing the semantic and syntactic relationships within the context and language present in the training dataset.

Textual content generation is performed based on the maximum conditional probability of the words in the training dataset, considering the context created by previously generated segments. For this reason, LLMs are considered probabilistic models.

These models offer several advantages. Firstly, they perform significantly better in text comprehension and generation in various domains. In addition, they have superior generalization capabilities, as they incorporate a broader knowledge of natural language.

However, these models also pose certain challenges, such as high computational resource consumption, the requirement for massive amounts of training data, and the risk of generating misleading answers. With regard to the latter, it should be noted that LLMs are based on probabilities calculated from the texts present in the training dataset. As a result, there is a considerable risk of generating plausible content, observing rules of syntax and semantics, yet containing untruthful content. This

13 ***The Large Language Model*** *is a generative model capable of generating synthetic content in large quantities and with high precision.*

risk has been referred to by experts as "hallucinations" (Beutel; Geerits; Kielstein, 2003).

Although generative pre-trained transformers have achieved impressive results in various natural language tasks, they lack a true understanding of text. They operate mainly based on statistical patterns learned during training and may occasionally generate responses that are nonsensical or grammatically incorrect.

It should be noted that language models are not exclusively based on ML, as there are models based on other theories, such as n-grams (Baeza-Yates; Ribeiro-Neto, 2013).

To conclude this chapter, another important concept for Generative AI is introduced: foundation models.

Foundation Model[14]

Foundation Model (FM), also referred to as the base model, is an extremely deep and complex neural network composed of billions of parameters. These parameters are adjusted during the training phase, which is performed on a massive amount of data, typically over a long period and in a distributed manner. After the training phase, FM can represent complex entities, such as human language, images, videos, and others (Fregly; Barth; Eigenbrode, 2023).

These models exhibit a relevant characteristic: the ability to perform tasks for which they were not explicitly trained. For example, an image analysis foundation model can be used to identify abnormalities in medical imaging, detect possible infrastructure failures to prioritize maintenance, among other uses (ITI, 2023). Therefore, the complexity of the FM allows it to be applied across a wide range of tasks.

It is common to find in the literature the association of foundation models solely with generative tasks. This occurs because foundation models are widely utilized in Natural Language Processing for text generation tasks. That being said, it is important to understand that these models are also

14 *A **foundation model** is a model with large parameters trained on a wide range of data sets in a self-supervised manner. The term refers to its importance and applicability in a wide range of domains (Gartner, c2024).*

used in image recognition and classification, speech-to-text conversion, and sentiment analysis, among other discriminative tasks (ITI, 2023). In summary, it is incorrect to assume that the terms "generative model" and "foundation model" are interchangeable.

‹ generative artificial intelligence and personal data protection ›

The Brazilian Data Protection Law - LGPD (Law No 13,709, of August 14, 2018) aims to protect the fundamental rights to privacy, freedom and the free development of personality (LGPD, Article 1, caput), while at the same time being grounded in economic and technological development and innovation (LGPD, Article 2, item V). Therefore, technological innovation must be pursued in harmony with the protection of personal data. In order to adequately protect *personal data*[15], it is necessary to understand how data is processed in Generative AI systems. The recent release of such systems for public use by society has demonstrated their rapid popularization and utilization for diverse purposes.

Generative AI systems exhibit the following fundamental characteristics: (i) the requirement for massive volumes of data for training; (ii) an inference capability that enables the generation of new data similar to the training data; and (iii) the adoption of a diverse set of computational techniques, such as transformer architectures for *Natural Language Processing*[16] (NLP) and ML algorithms, exemplified by generative adversarial networks for generating visual content.

Such characteristics directly affect the *personal data processing*[17] and the principles governing the LGPD. The requirement for massive volumes of data can result in the processing of both personal and non personal data. The coexistence of these two types of data elevates the risks related to personal data protection, as it increases the probability that personal

15 ***Personal data:*** *information relating to an identified or identifiable natural person.*

16 ***Natural Language Processing:*** *is an interdisciplinary area that involves the field of computer science dedicated to the interaction between computers and human language, through programs capable of processing, analyzing, interpreting, and generating natural language data. This includes text summarization, machine translation, speech recognition, sentiment analysis, text and speech generation, among others. Natural language is any human language, which can be expressed in text, speech, sign language etc.*

17 ***Personal data processing:*** *any operation carried out with personal data, such as collection, production, reception, >*

data will be processed without proper safeguards and that the necessity principle will not be met. Furthermore, the capability to generate new data, or synthetic content, poses a risk to personal data protection, since the synthetic content generated may be indistinguishable from personal data and may relate to an identified or identifiable natural person, as well as being erroneously associated with real individuals. Finally, the set of computational techniques results in complex and opaque methodologies, meaning that the low level of transparency is not necessarily due to the nature of the technique used, but rather to the difficulty of interpreting its operations and understanding the processes used to obtain the results.

Thus, *data subjects*[18] and *controllers*[19] or *processors*[20] are faced with AI systems that pose significant challenges for 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 compliance with principles such as transparency. These conditions require close attention to the principles and rules established by the LGPD and pose challenges related to the risk of violations of fundamental rights.

The following section examines the relationship between generative AI and different personal data processing operations. Furthermore, brief considerations regarding generative AI and certain LGPD principles will be presented at the end of this chapter. To this end, the following reference framework was adopted: CNIL (c2023), Glenster and Gilbert (2023), ABNT (2023), AEPD (2023), OPC (2023), Solove (2024), Rana *et al.* (2022), and Teffé (2017).

The relationship between the personal data processing and Generative AI systems

In order to highlight the relationship between the processing of personal data and Generative AI systems, this study outlines four (4) sets of elements that comprise personal data processing: (i) collection and storage, (ii) processing, (iii) sharing, and (iv) deletion.

> classification, use, access, reproduction, transmission, distribution, processing, archiving, storage, deletion, evaluation or control of information, modification, communication, transfer, dissemination, or extraction.

18 ***Data subjects:*** *natural person to whom the personal data being processed refer.*

19 ***Controller:*** *a natural or legal person, governed by public or private law, who is responsible for the decisions regarding the processing of personal data.*

20 ***Processor:*** *natural or legal person, governed by public or private law, who processes personal data on behalf of the controller.*

i) Data collection and storage for training

The development of Generative AI models involves several stages, and one of the first steps is the collection of data that is subsequently stored and used to train the model.

One method for collecting data to build massive datasets for training and testing Generative AI systems is through the technique of data scraping (web scraping). This technique uses web-browsing programs that collect, extract, and/or copy data created and made available by web users or third parties. The collected data may include any information available on the web, such as first and last names, addresses, e-mail addresses, videos, audios, images, comments, opinions, preferences, among other data and identifiers, available on different websites or databases.

Depending on the automation method used, data scraping may select specific data, for instance, collecting only opinions, and can even define regular time intervals to perform new collections. In other cases, scraping can operate on a significant scale and traverse billions of webpages. The dynamic nature and rapid availability of new data on the web enable the development of different systems for data scraping and aggregation *(data aggregators)*[21].

Common Crawl stands out as an example of a massive dataset available and utilized for training *AI models*[22]. This is a non-profit organization that regularly crawls and collects web data using *crawlers*, that is, automated programs, with the aim of generating repositories. It constitutes a centralized infrastructure that aggregates data from different sources, which are stored and made available free of charge and openly to anyone who wants to carry out research and develop innovations that require massive volumes, including in the field of AI[23].

The comprehensive nature of massive data collection enables repositories such as those maintained by *Common Crawl* to offer a wide and diverse source of data obtained through data scraping and aggregation. For the training of AI models, developers can mix data scraping carried out directly by the developer with sources from another organization such as *Common Crawl*, as well as with data from a variety of addition sources,

21 ***Data aggregation:*** *a technique which groups data collected from a wide variety of sources into a centralized repository for access and processing through computer systems.*

22 ***AI models:*** *mathematical and logical representation of a system, entity, phenomenon, process, or data. The model enables AI to process, model, adapt, interpret, and generate data to interact with the world, with the ability to generalize learning for different contexts. Models can be divided into different types according to their function or application, for example, machine learning models, rule-based models, artificial neural network models, among others.*

23 *More information can be found at: https://commoncrawl.org/*

including books, scientific articles and publications, dissertations and theses, audio and video transcripts, tables, codes, legislation, among others.

The operation of large-scale data scraping and aggregation increases the risks associated with the potential inclusion of personal data. The wide scope of data that can be collected for model development raises similar concerns.

The absence of adequate pre-processing steps for the deletion or *anonymization*[24] of personal data means that there are significant risks related to the improper processing of personal data. These risks are exacerbated in cases involving the processing of *sensitive personal data*[25] and the data of children and adolescents.

It should be noted that the content of public or publicly accessible websites remains subject to the LGPD, since companies that make such information available have obligations regarding the protection of personal data on their platforms. Likewise, developers and companies that perform web scraping must ensure compliance with personal data protection frameworks.

Processing personal data without the knowledge of the data subjects limits their control over their personal data. Such loss of control may occur even after the data subjects have deleted the data on the web, due to the possibility of a prior scrape and its storage in repositories.

Data scraping operations must indicate the legal basis for the processing of personal data. In this sense, companies carrying out this activity need to consider one of the legal hypotheses present in Articles 7 and 11 of the LGPD. In addition, pursuant to Article 7, Paragraphs 3 and 4, the law expressly provides for the application of data protection principles in cases where personal data are made public by the data subject or are publicly accessible. Therefore, the use of scraped personal data must pay close attention, above all, to the principles of good faith, purpose, adequacy, and necessity.

Therefore, mechanisms are required to ensure adequate transparency in the data collection and storage stages for the formation of massive datasets.

24 ***Anonymization:*** *the use of reasonable technical means, available at the time of processing, by which a piece of data loses the possibility of direct or indirect association with an individual.*

25 ***Sensitive personal data:*** *personal data on racial or ethnic origin, religious conviction, political opinion, membership of a trade union or religious, philosophical or political organization, data relating to health or sex life, genetic or biometric data, when linked to a natural person.*

ii) Processing

Data processing spans multiple stages in the *Generative AI lifecycle*[26]. Data processing activities begin in the phase prior to training the model, namely during the creation of training and testing datasets, and continue throughout the entire Generative AI lifecycle.

The utilization of data from repositories originating from web scraping processes, as well as the combination of data scraped directly by developers with data from other sources, may result in the use and reuse of data for training and refining different AI models. Where personal data is present, this may lead to continuous, unrestricted, and unlimited processing of personal data across different Generative AI systems.

The ability of Generative AI systems to generate new synthetic content goes beyond basic data processing, encompassing learning and modeling to generate new representations from training data.

During training, model parameters are assigned their specific values, which represent *weights*[27] associated with the model's capability to generate, for example, accurate natural language. The weights reflect learned patterns and relationships, that is, how the model responds to and learns from the training data.

Models do not process or store specific information, as their mathematical nature obfuscates the underlying data. In this sense, the existence of personal data in the datasets can be hidden by means of mathematical processes during training, such that personal data may not be directly identified in the model. Although it should be noted that recent studies have discussed the possibility of natural persons being identifiable due to vulnerabilities to model inversion and membership inference attacks (Veale, Binns; Edwards, 2018).

However, Generative AI systems allow users to interact in natural language with the model trained to generate answers. Thus, depending on the form of interaction, the instructions provided, and the context supplied by the user through the *prompt*[28], personal data may be generated as an output by Large Language Models (LLMs).

26 ***Generative AI lifecycle:*** *a model that describes the evolution and stages of an AI system, from the start of its development to its decommissioning. The stages are not sequential and can often occur iteratively, such as risk management, corrections, refining the model, implementing improvements, and updating the system. The decision to decommission an AI system can occur at any time during the operation and monitoring phase.*

27 ***Weights:*** *internal variable in a model that affects the way outputs or results are calculated. In the case of Generative AI, weights determine the relative importance of different inputs for calculating outputs. To do this, the weights are iteratively adjusted during the model's training process to capture patterns and characteristics of the training data.*

28 ***Prompt:*** *can be translated as input or command. In computing, a prompt is a message or a command line in a user interface. In the context of Generative AI, the prompt is a text input used to give instructions to the AI model on what to do or what question it should answer.*

The *content*[29], although synthetic, meaning it is generated by the model, presents narratives that may result in the production of false or inaccurate information about a real person. This possibility poses risks to data protection and to the free development of personality, especially regarding the data subject's right to their image.

At this point, the issue does not concern only a person's physical appearance or portrait, that is, their image-portrait as an external expression of the human person, but also their image-attribute, which is related to a set of characteristics that can be associated with a person in their representation in the social environment, such as honor, reputation, dignity, or prestige.

For example, in an interaction involving a prompt for résumé screening, the system can generate synthetic content that presents unfounded and untrue facts about the candidate's personal life, misinterpretations of opinions, false information, or any other content that can disseminate misinformation or harm an individual's reputation. Another example is the case of singer Taylor Swift, who was the target of viral dissemination of false sexual images generated by *deepfakes*[30].

Daniel Solove (2024) refers to this untruthful content generated by generative AI as "malevolent material" due to its ability to amplify fraud and other scams. He warns, however, that the Generative AI operator does not always generate these materials intentionally, due to the phenomenon of *hallucination*[31]. Intent may be a relevant element in liability regimes based on a subjective conception of fault; however, it is less relevant in strict liability regimes or those grounded in a normative conception of fault.

Accordingly, a key challenge lies in defining who should be held responsible for the generation of hallucinations concerning natural persons that produce a harmful effect. Another challenge concerns compliance with the LGPD, since Generative AI systems can generate personal data without having been specifically trained for this purpose.

29 ***Synthetic content:*** *a set of information made up of synthetic data. In this section, we have chosen to use this expression to emphasize that the content generated may include personal data. It is also important to bear in mind that the term "synthetic data" is used both in the field of computing and in privacy and data protection. In computing, synthetic data can be used in the development and testing of AI systems when real data is not available in the quantities required, does not exist or cannot be processed, and for balancing databases. In the area of privacy and data protection, the use of synthetic data is often associated with Privacy Enhancing Technologies (PETs). In this sense, AEDP emphasizes that synthetic data will be a PET technique if it is used to generate sets of non-personal data with the same utility as the personal data. It is important to emphasize, however, that "synthetic content" is not generated for the purpose of anonymization.*

30 ***Deepfakes:*** *the term "deepfake" is derived from the terms "deep learning" and "fake". The term describes manipulated realistic content such as photos and videos generated by deep learning.*

31 ***Hallucination:*** *the term hallucination in Large Language Models*

iii) Sharing

Due to the broad scope of the concept of sharing in the processing of personal data, which can be approached from different perspectives, this study divided sharing into three stages: (1) data sharing by the user of the Generative AI systems, who may or may not be the data subject; (2) the sharing by third parties of results obtained through prompt interactions in Generative AI systems that contain personal data; and (3) the sharing of pre-trained models containing personal data.

(LLMs) is characterized by generated content that is not representative, truthful or does not make sense in relation to the source provided, for example, due to errors in coding and decoding between text and representations. However, it should be noted that artificial hallucination is not a new phenomenon (Beutel; Geerits; Kielstein, 2003).

1. *Data sharing by users of the Generative AI systems, who may or may not be the data subjects*

First, it is worth considering on the sharing of new personal data by users who interact with these systems (whether acting as data subjects, controllers, or processors) through command prompts.

The prompt functionality of current Generative AI enables the sharing of a wide range of data to generate answers. As these systems have evolved, prompt functionality has been enhanced to support attachments in different formats. As a result, users can provide instructions and attach documents that directly entail data sharing and processing.

The instructions provided by users may include a variety of information, such as excerpts from documents, text messages, emails, comments, and opinions available on different platforms, details of personal experiences, academic research, purchase history, interactions with customers, medical records, questions, and reports on medical procedures, among others, which may contain personal data and sensitive personal data.

In addition, users can have the option to upload entire documents to broaden interaction and data-processing capacity. Thus, confidential business documents, medical prescriptions and reports, public documents such as deeds and powers of attorney, minutes of meetings, spreadsheets, figures, among others, can be attached to

receive analyses, interpretations, or any other output that the user considers pertinent. In this way, the information made available to the prompt is used to learn about the context.

A relevant aspect of LLMs is the generation of personalized responses. The information provided in the prompt is used to model the answers within a given context, such that the context of the previous answer can be used to answer a subsequent question on the same subject.

In many cases, the processing agent and the data subject may be unaware of the risks involved in this sharing of information or may trust the system due to the benefits provided by the results or assistance received. Furthermore, if a natural person using the AI system shares personal data of other data subjects with the Generative AI system, that user may, depending on the context, be considered a processing agent.

In this respect, systems should be developed to protect users' privacy in interactions that may involve sharing personal data in the prompt. Another relevant aspect is the absence of clear and easily accessible information on the processing of personal data made available in the prompt.

Shifting responsibility for the protection of personal data onto the user in order to guarantee the proper use of Generative AI systems that adopt LLMs does not seem to be enough to deal with the consequences that currently exist from sharing personal data via the command prompt.

2. *Third-party sharing of prompt interaction outputs containing personal data within Generative AI systems*

Generative AI systems can allow personal data to be shared with third parties when that data constitutes the synthetic content generated by these systems.

In this case, the observations mentioned in the previous section (Processing) must be carefully adhered to, especially considering the risk of the shared data being reused for secondary purposes that the developer of the Generative AI system is unlikely to be able to control.

Establishing a clear chain of accountability between the different actors involved in this sharing of personal data becomes a relevant, albeit challenging, element for ensuring compliance with the LGPD.

3. Sharing of pre-trained models containing personal data

Since pre-trained models can be considered a reflection of the dataset used during training, the widespread adoption of *APIs*[32] that utilize foundation models such as pre-trained LLMs brings a new challenge. The sharing of models also tends to involve the data that are mathematically embedded within them.

This type of sharing enables the development of independent applications that *fine-tune*[33] or refine the foundation model by training it with a set of data specific to the intended domain.

By linking the fine-tuning of the foundation model with the possibility of using the prompt interaction outputs to generate training bases for refinement, the existence of personal data triggers a continuous cycle of processing, as will be described in the next topic (Deletion).

The sharing of foundation models that have been trained on personal data, as well as the use of such data for refinement, may entail distinct data protection risks, depending on the intended purpose.

32 ***APIs:*** *acronym for Application Programming Interface; a way of allowing, through software programs, interaction between different applications and extracting data.*

33 ***Fine-tuning:*** *technique that can be used in Machine Learning to adjust a model that has already been trained on a large set of data for use in a specific domain. Tuning takes place by re-training with a more restricted dataset.*

iv) Deletion

The deletion stage refers to the phase in which personal data, or datasets stored in databases, are erased upon the termination of their processing.

The definition of the end of the processing of personal data in Generative AI systems must account for three new relevant elements: the generation of synthetic content; prompt interactions that allow for the sharing of new data; and the continuous refinement of the model.

The integration of these elements into a single Generative AI system may involve personal data. This new technological context may result in the possibility of continuous processing of personal data and demand new approaches.

In this way, there is a challenge in delimiting the end of the processing period, as well as in determining whether the purpose or need for processing has been fulfilled. Additional difficulties relate to the effective withdrawal of consent of the data subject in Generative AI systems (should this legal basis be applied).

It is therefore important to observe the principle of accountability (Article 6, item X) for all actors in the supply chain of Generative AI systems, for as long as such personal data has not been definitively deleted.

In short, the entire lifecycle of personal data processing and the use of elements of Generative Artificial Intelligence must be compatible with the rights and freedoms of individuals, ensuring that the rights and principles that guide the LGPD are observed.

Generative AI and the principles of the LGPD

Regarding the principles, a few points can be made. The *principle of transparency* requires clear, precise, and easily accessible information to data subjects. It is common for data subjects not to be informed about the scraping of their data on the web, the inclusion of their personal data in model training datasets, or the possibility that prompt interactions may involve the sharing of their personal data or that of third parties. Consequently, there is a frequent lack of detailed technical and non-technical documentation on the processing of personal data in different Generative AI systems.

The existence of detailed technical documentation would serve as a baseline for verifying compliance with personal data protection requirements and with respect to the data sources used, aggregated, and filtered, in order to highlight the techniques or practices adopted to prevent the use of personal data. Likewise, the production of appropriate documentation could

help to monitor Generative AI systems throughout their lifecycle, to identify improvements, and enable the exercise of rights where personal data are involved. Such documentation could mitigate the risks related to applications involving the processing of personal data and guarantee transparency.

In turn, the *principle of necessity* presents an additional challenge related to the use of large datasets in modern Generative AI systems, particularly regarding compliance with the requirement to limit processing to what is strictly necessary to achieve the intended purpose. The principle does not indicate a ban on the training of Generative AI systems with large volumes of data, but demands careful consideration and care before training, to avoid the existence of unnecessary personal data in the training datasets, as well as being inserted later through the prompt or attachments.

Similarly, although the other principles of the LGPD also do not prohibit growth and innovation in the field of Generative AI, they introduce considerations that must be observed for the development and responsible use of these technologies. After all, it is necessary to guarantee the responsible development and full progress of Generative Artificial Intelligence across different sectors, in conjunction with respect for privacy and personal data protection throughout its entire lifecycle.

‹ generative artificial intelligence in the Brazilian context ›

In the Brazilian context, several initiatives involving the adoption of Generative AI systems can be observed. For illustrative purposes, this study presents three use cases: (i) a public institution; (ii) the healthcare sector; (iii) the banking sector. However, it is important to note that there are numerous other contexts of use beyond those highlighted here, which have not been included due to the limited scope of this document.

i) Public institution

In 2023, the Brazilian Federal Court of Accounts (TCU) adopted a customized writing assistant model based on Natural Language Processing and Artificial Intelligence, called ChatTCU (TCU, 2023).

ChatTCU[34] was developed following studies by a Working Group (WG) tasked with evaluating the risks and opportunities of using AI tools. The initial objective of the assistant is to optimize time and to support the Court's teams in content drafting, adaptations to plain language, translations, and analyses of external control actions.

The Court states that ChatTCU will receive updates as new features are developed, and, in the future, the system is expected to have full access to the TCU database to provide relevant and accurate information on specific cases managed by the Court.

One of the system's features is that conversations are protected by contractual safeguards that guarantee the confidentiality of the information provided by users. In addition, the existence of an internally used AI use offers greater security when compared to an AI for public use (such as ChatGPT).

34 *More information is available in Portuguese at: https://portal.tcu.gov.br/imprensa/noticias/tcu-adota-modelo-personalizado-de-assistente-de-redacao-baseado-em-inteligencia-artificial.htm.*

ii) Healthcare sector

In 2023, the Bill & Melinda Gates Foundation selected Munai, a *healthtech* company based in Curitiba to develop an Artificial Intelligence project based on Large Language Models, such as ChatGPT (PMC, 2023).

The project aims to withstand the effects related to the fight against antimicrobial resistance (AMR), which is the ability of microbes or bacteria not to succumb to the effects of medications such as antibiotics, the unnecessary use of antibiotics, and the lack of adherence to the Antimicrobial Stewardship Program (ASP).

As part of the project, a virtual assistant will be developed to automate hospital protocols, provide clinical decision support, and improve accessibility, interpretation, and application of such protocols. In practice, the

solution aims to assist physicians and prescribers in making rational use of antibiotics, given that currently about half of antimicrobial prescriptions are unnecessary. The project may be deployed in Brazilian hospitals as well as in international healthcare institutions.

The company was one of 51 organizations selected by the foundation, which aims to promote solutions with a social impact on global issues using Large Language Models. Munai already operates in the healthcare management sector, utilizing data and AI.

iii) Banking sector

Banco do Brasil developed an assistant that uses Generative AI based on the ChatGPT Large Language Model to deliver responses tailored to the institution's business context, supporting frontline customer service personnel (BB, 2023).

The technique employed combines interactions with the Language Model via prompt engineering with information retrieval through semantic indexing, which uses embedding spaces to capture and organize data meanings. The system is currently in its deployment phase, and the topics covered include issues related to mortgages, auto loans, renewable energy, sustainability actions, and credit cards.

Employees will be able to query the intelligent assistant when customers contact the bank regarding any of these topics. The generated outputs will be used both to help the employee provide information and to facilitate customer decision-making.

The model will be updated through ongoing use and training to deliver increasingly accurate, refined, and personalized responses based on the volume of interactions surrounding the customers' most frequently asked questions.

‹ prospects ›

Generative models have demonstrated significant potential for generating increasingly realistic content. This potential unfolds many opportunities for innovation and technological development, but it will undoubtedly introduce major data protection and privacy concerns. These issues will present ongoing challenges that society must address.

A recent breakthrough in the field that warrants attention is associated with the *Reinforcement Learning*[35] paradigm: improvements in the algorithm known as *Q-Learning*[36] have the potential to reshape the generative AI landscape. The implications of these advancements are vast, paving the way for applications with the potential to surpass the capabilities of generative models. Consequently, there is an expectation with Q-Learning that the current limitation of generative models in solving mathematical problems where only one correct answer exists, as opposed to content generation and translation, where there is a wide variety of right answers, will be overcome. This would be done by aligning the deterministic aspects of traditional AI algorithms with the creative and generative potential of LLMs (McIntosh *et al.*, 2023).

In the commercial technology market, Google recently launched the Gemini product (Google, c2024). Gemini is a *multimodal language model*[37] with the ability to process data of distinct types, such as text, images, or code. In technical terms, Gemini is built using a multimodal transformer architecture with billions of parameters, enabling the model to understand the meaning of information in depth.

The scarcity of open-source detailed models in the field of AI restricts opportunities for more comprehensive studies aimed at understanding impacts and exploring their innovative potential. To bridge this gap, models are gradually emerging, such as the OLMo *(Open Language Model)* project, developed by the Allen Institute for AI (AI2). OLMo provides full access to the LLM, training data, source code, model weights, and documentation (AI2, c2024).

35 ***Reinforcement Learning*** *is a machine learning model where the model is trained only in terms of positive reinforcement (reward) and negative reinforcement (penalty). During problem solving, the model performs actions so that the overall reward is maximized and, at the same time, the punishments are minimized (Gartner, c2024).*

36 ***Q-Learning*** *is an algorithm from the reinforcement learning paradigm that aims to find the best action to take in a given circumstance.*

37 ***A multimodal language model*** *can process and generate data from different modalities, such as text, images, sounds, and code.*

The model has undergone extensive training on 3 trillion tokens, which ensures robust performance in text generation and reading comprehension. The possibility of accessing the project's open-source code and its documentation is an important resource that can promote collaboration, transparency, and new developments in the field of Generative AI.

The challenges of LLM reliability and accuracy are seeing advances through the use of techniques such as CRAG *(Contextual Retrieval Augmented Generation)*, which seek to enrich the process of generating synthetic content and mitigate hallucinations. Currently, these challenges persist because the accuracy of the generated texts cannot be guaranteed solely by the model's acquired parametric knowledge (Yan *et al*., 2024). CRAG introduces a retrieval evaluator that performs contextual evaluation of external documents, focusing on key information retrieval actions. This process aims to ensure that reliable information is incorporated into the synthetic content generated (Yan *et al.*, 2024).

Advances in the field are constant, and the current steps demonstrate a push toward *General Purpose Artificial Intelligence* (GPAI)[38]. GPAI is characterized as a type of AI with general problem-solving capabilities. Although GPAI is on the horizon of scientific advancement, it represents a major step towards creating systems that are highly adaptable to a variety of scenarios. However, there is still a long way to go to achieve GPAI. It is important to note that the concept of GPAI should not be confused with that of *Strong Artificial Intelligence*[39], which is associated with issues of sentience and consciousness.

38 ***General Purpose Artificial Intelligence (General Purpose AI)*** *is the form of artificial intelligence that can understand, learn and apply knowledge to a wide range of tasks and domains. General Purpose Artificial Intelligence can be applied to a much broader set of use cases and incorporates cognitive flexibility, adaptability, and general problem-solving skills (Gartner, c2024). The European Parliament considers ChatGPT to be an example of general--purpose AI (European Parliament, 2023).*

39 ***Strong Artificial Intelligence*** *is the branch of artificial intelligence that considers the hypothetical capability of a machine actually thinking, rather than simulating thinking (Russel; Norvig, 2013).*

‹ final considerations ›

Data protection in the context of Generative AI must be approached from an ethical, legal, and sociotechnical perspective. Technological innovation in the field of Generative AI may introduce new risks or amplify existing ones.

The broad range of applications of Generative AI could have diverse impacts on society, such as distancing individuals from what is real, produced, thought, and certified by humans. This could affect the perception of the world and entail risks that are still unknown to society. The trust placed in the outputs generated by these models requires further examination, making it necessary to evaluate the positive and negative aspects involved in their deployment.

Data is a fundamental element for developing various AI systems; however, in the field of Generative AI, systems also generate data and enable new forms of information sharing. Data becomes integral to the entire lifecycle of the system, and its continuous flow is renewed with every generated output or shared file that may contain personal data.

In Generative AI, concerns regarding the presence of personal data within the trained model extend to model refinement and the possibility of new developments utilizing trained models that may involve personal data processing.

The variety of future applications already under development across sectors such as healthcare, public institutions, and banking demonstrates their societal impact and direct relationship with personal data protection. These issues warrant further debate, as economic and technological development, innovation, respect for privacy, and the foundational pillars of the LGPD are essential for the ethical and responsible growth of Artificial Intelligence and Generative AI across different domains and tasks.

At this stage, given that this is a preliminary study, only a limited number of system risks and vulnerabilities related to personal data processing

have been identified; these should not be exhaustive. Consequently, continued studies in this field are necessary to guarantee adequate protection for data subjects throughout the evolutionary lifecycle of these technologies and to promote the development of ethical and responsible AI systems.

‹ references ›

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). **ABNT NBR ISO/IEC 22989: 2023:** Tecnologia da informação — Inteligência artificial — Conceitos de inteligência artificial e terminologia. 1 ed. Rio de Janeiro, 2023.

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS (AEPD). Synthetic data and data protection. **AEPD**, November 2, 2023. Available at: https://www.aepd.es/en/prensa-y-comunicacion/blog/synthetic-data-and-data-protection. Accessed on: January 11, 2024.

ALLEN INSTITUTE FOR IA (AI2). OLMo. **AI2**, c2024. Available at: https://allenai. org/olmo. Accessed on: February 05, 2024.

BAEZA-YATES, Ricardo.; RIBEIRO-NETO, Berthier. **Recuperação de Informação:** Conceitos e Tecnologia das Máquinas de Busca. 2. Ed. Bookman, 2013.

BANCO DO BRASIL (BB). BB usa tecnologia generativa para apoiar atendimento. **BB**, Jun. 27, 2023. Available at: https://www.bb.com.br/pbb/pagina-inicial/imprensa/n/67461/bb-usa-tecnologia-generativa-para-apoiar-atendimento#/. Accessed on: February 1, 2024.

BEUTEL, Gernot; GEERTIS, Eline; KIELSTEIN Jan T. Artificial Hallucination: GPT on LSD? **Critical care**, v. 27, n. 1, April 2023.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados. Brasília, DF: Presidência da República, 2018. Available at: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm. Accessed on: July 11, 2023.

COMMISSION NATIONALE DE L'INFORMATIQUE ET DES LIBERTÉS (CNIL). AI how-to sheets. **CNIL**, c2023. Available at: https://www.cnil.fr/fr/ai-how-to-sheets. Accessed on: February 2, 2024.

CRISTIANINI, Nello; SHAWE-TAYLOR, John. **An Introduction to support Vector Machines and Other Kernel-based Learning Methods**. London: Cambridge University Press, 2000.

DARTMOUTH. Artificial Intelligence Coined at Dartmouth. **Dartmouth**, c2023. Available at: https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth. Accessed on: July 07, 2023.

EUROPEAN PARLIAMENT. General-Purpose artificial intelligence. **European Parliamentary Research Service,** March, 2023. Available at: https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/745708/EPRS_ ATA(2023)745708_EN.pdf. Accessed on: February 6, 2024.

FIX, Eevelyn; HODGES, J. L. **Discriminatory Analysis Nonparametric Discrimination:** Consistency Properties. USAF School of Aviation Medicine, 1951.

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

FREGLY, Chris; BARTH, Antje; EIGENBRODE, Shelbee. **Generative AI on AWS:** Building Context-Aware Multimodal Reasoning Applications. Sebastopol, CA: O' Reilly, 2023.

GARTNER. Gartner Glossary. **Gartner**, c2024. Available at: https://www.gartner.com/en/information technology/glossary/foundation-models. Accessed on: February 2, 2024.

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

GOODFELLOW, Ian J. *et al.* Generative Adversarial Networks. **Communications of the ACM**, New York, v. 63, n.11, 2014.

GOOGLE. Gemini. **Google DeepMind**, c2024. Available at: https://deepmind.google/technologies/gemini/#introduction. Accessed on: January 2, 2024.

HARRIS, Laurie; ZHU, Ling. **Generative Artificial Intelligence and Data Privacy:** A primer. [S. l.]: Congressional Research Service, May 23, 2023. Available at: https://www.congress.gov/crs-product/R47569. Accessed on: June 11, 2026.

HAYKIN, Simon. **Neural Networks:** A Comprehensive Foundation. 2. ed. Upper Saddle Rive: Prentice-Hall, 1999.

INFORMATION TECHNOLOGY INDUSTRY COUNCIL (ITI). Understanding AI Foundation Models & The AI Value Chain: ITI's Comprehensive Policy Guide. **ITI**, Aug., 2023. 11 p. Available at: https://www.itic.org/documents/artificial-intelligence/ITI_AIPolicyPrinciples_080323.pdf. Accesses on: 24 abr. 2026.

CHOLLET, François. Grad-CAM class activation visualization. **Keras**, April 26, 2020, modified in March 7, 2021. Available at: https://keras.io/examples/vision/grad_cam/#lets-testdrive-it. Accessed on: January 9, 2024.

MCINTOSH, Timothy; SUSNAJAK, Teo; LIU, Tong; WATTERS, Paul; HALGAMUGE, Malka. From Google to OpenAI Q* (Q-Star): A Survey of Reshaping the Generative Artificial Intelligence (AI) Research Landscape. **Journal of Latex Class Files,** v. 1, n. 1, dec. 2023.

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

MONTGOMERY, Blake. Taylor Swift AI Images Prompt US Bill to Tackle Nonconsensual, Sexual Deepfakes. **The Guardian** (Jan. 30, 2024). Available at: https://www.theguardian.com/technology/2024/jan/30/taylor-swift-ai-deepfake-nonconsensual-sexual-images-bill?ref=biztoc.com. Accessed on: February 2, 2024.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OCDE). OECD Ai Principles overview. **OCDE**, c2024. Available at: https://oecd.ai/en/ai-principles. Accessed on: April 17, 2024.

OFFICE OF THE PRIVACY COMMISSIONER OF CANADA (OPC). Joint statement on data scraping and protection of privacy. **OPC**, August 24, 2023. Available at: https://www.priv.gc.ca/en/opc-news/speeches-and-statements/2023/js-dc_20230824/. Accessed on: February 2, 2024.

PREFEITURA MUNICIPAL DE CURITIBA (PMC). Startup de Curitiba é selecionada pela Fundação de Bill Gates para desenvolver projeto de IA. **Prefeitura Municipal de Curitiba**, August 8, 2023. Available at: https://www.curitiba.pr.gov.br/noticias/startup-de-curitiba-e-selecionada-pela-fundacao-de-bill-gates-para-desenvolver-projeto-de-ia/69799. Accessed on: January 31, 2024.

QUINLAN, J. Ross. Discovering rules by induction from large collections of examples. *In*: MICHIE, Donald (Ed.) **Expert Systems in the Microelectronic.** Edinburgh: Edinburgh University Press, 1979. p. 168-201.

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

RUSSEL, Stuart.; NORVIG, Peter. **Inteligência Artificial.** Tradução de Regina Célia Simille. Rio de Janeiro: Elsevier, 2013.

SOLOVE, Daniel J. Artificial Intelligence and Privacy. **GWU Legal Studies Research Paper**, n. 36, 2024. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4713111. Accessed on: February 2, 2024.

TRIBUNAL DE CONTAS DA UNIÃO (TCU). TCU adota modelo personalizado de assistente de redação baseado em inteligência artificial. **TCU,** June 6, 2023. Available at: https://portal.tcu.gov.br/imprensa/noticias/tcu-adota-modelo-personalizado-de-assistente-de-redacao-baseado-em-inteligencia-artificial.htm. Accessed on: January 16, 2024.

TEFFÉ, Chiara S. Considerações sobre a proteção do direito à imagem na internet. **Revista de Informação Legislativa**, v. 54, n. 213, p. 173-198, Jan./Mar. 2017.

TONG, Xiaochu *et al.* Generative Models for De Novo Drug Design. **Journal of medicinal chemistry**, v. 64, n. 19, 2021.

USZKOREIT, Jakob. Transformer: A Novel Neural Network Architecture for Language Understanding. **Google Research**, August 31, 2017. Available at: https://research.google/blog/transformer-a-novel-neural-network-architecture-for-language-understanding/. Accessed on: May 4, 2026

VASWANI, Ashish *et al.* Attention Is All You Need. *In*: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEM (NIPS), 31. **Proceedings** […]. Long Beach, CA, 2017.

VEALE, Michael; BINNS, Reuben; EDWARDS, Lilian. Algorithms that remember: Model inversion attacks and data protection law. **Philosophical Transactions of the Royal Society A**, v. 376, n. 2133, 2018.

YAN, Shi-Qi *et al.* Corrective Retrieval Augmented Generation. *In*: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 30. **Proceedings** […]. Singapore, 2025. arXiv, 7. Out. 2024. https://doi.org/10.48550/arXiv.2401.15884.

www.gov.br/anpd

ANPD