Response by the Centre for Information Policy Leadership (CIPL) to Brazil's ANPD Consultation on Artificial Intelligence and Automated Decisions (Art. 20)

Submitted January 24, 2025

The Centre for Information Policy Leadership (CIPL) welcomes the opportunity to respond to Brazil's ANPD's Consultation on Artificial Intelligence and Automated Decisions (Art. 20 of LGPD). CIPL commends the ANPD's efforts to encourage the responsible development and deployment of AI. For more than 20 years, CIPL has been a thought leader on organisational accountability and a risk-based approach as key building blocks of smart regulation, responsible governance, and use of data, as well as accountable development and deployment of AI.[1] CIPL's "Ten Recommendations for Global Regulation" proposes a layered, three-tiered approach to AI regulation that would protect fundamental human rights and minimise the potential risks of harm to both individuals and society, while enabling the responsible development and deployment of AI.[2] Our benchmarking "report, "Building Accountable AI Programs: Mapping Emerging Best Practices to the CIPL Accountability Framework", outlines best practices and case studies on how 20 leading organisations are responsibly developing and deploying AI through the lens of CIPL's Accountability Framework.[3] CIPL's most recent discussion paper, "Applying Data Protection Principles to Generative AI: Practical Approaches for Organizations and Regulators", considers key privacy and data protection concepts and explores how they can be effectively applied to the development and deployment of generative AI models and systems.[4]

For reference, LGPD **Article 20** states as follows:

Article 20. The data subject is entitled to request the review of decisions made solely based on automated processing of personal data that affect his/her interests, including decisions intended to

---

[1] The Centre for Information Policy Leadership (CIPL) is a global privacy and data policy think tank in the law firm of Hunton Andrews Kurth LLP and is financially supported by the law firm and 85+ member companies that are leaders in key sectors of the global economy. CIPL's mission is to engage in thought leadership and develop best practices to ensure the responsible and beneficial use of data in the modern information age. CIPL's work facilitates constructive engagement between business leaders, data governance and security professionals, regulators, and policymakers around the world. For more information, please see CIPL's website at https://www.informationpolicycentre.com/. Nothing in this document should be construed as representing the views of any individual CIPL member company or of the law firm Hunton Andrews Kurth LLP. This document is not designed to be and should not be taken as legal advice.

[2] CIPL, "Ten Recommendations for Global AI Regulation", October 2023, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ten_recommendations_global_ai_regulation_oct2023.pdf.

[3] CIPL, "Building Accountable AI Programs: Mapping Emerging Best Practices to the CIPL Accountability Framework", February 2024, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_building_accountable_ai_programs_23_feb_2024.pdf

[4] CIPL, "Applying Data Protection Principles to Generative AI: Practical Approaches for Organizations and Regulators", December 2024, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_applying_data_protection_principles_genai_dec24.pdf

define his/her personal, professional, consumer and credit profile or aspects of his/her personality. (New wording given by Law No. 13,853/2019)

> **Paragraph 1.** The controller shall provide, upon request, clear and adequate information on the criteria and procedures used for an automated decision, complying with trade and industrial secrets.
>
> **Paragraph 2.** In the event of failure to offer the information set forth in paragraph 1 of this article based on the compliance with trade and industrial secrets, the national authority may carry out an audit to verify discriminatory aspects in the automated processing of personal data.

BLOCK 1 – PRINCIPLES OF THE LGPD

The General Law on the Protection of Personal Data (LGPD) is a regulatory framework that governs the processing of personal data, including digital data, by natural persons or legal entities governed by public or private law, with the aim of protecting the fundamental rights of freedom and privacy and the free development of the personality of natural persons.

In this sense, the processing of personal data must comply with the principles, rights and guarantees set out in the LGPD, regardless of the medium, physical or digital, or the technology used, such as Artificial Intelligence (AI) systems.

The development and use of AI systems must be guided by the principles that guide the processing of personal data, including the following:

(i) purpose, which limits the use of data to legitimate, specific, explicit and informed purposes, without the possibility of further processing in a manner incompatible with those purposes;

(ii) necessity, which requires that only strictly necessary data be used to achieve the purposes of processing;

(iii) data quality, which guarantees data subjects the accuracy, clarity, relevance and updating of the data, in accordance with the need and for the fulfillment of the purpose of its processing;

(iv) transparency, which requires the provision of clear, precise and easily accessible information about the processing and the respective processing agents; and

(v) non-discrimination, which prohibits the processing of personal data for unlawful or abusive discriminatory purposes.

(vi) responsibility and accountability, in which the agent must demonstrate the adoption of effective measures capable of proving compliance with personal data protection rules, including the effectiveness of these measures.

These principles are essential for the development and responsible use of AI systems that respect the rights of data subjects, avoiding excessive or inappropriate use of personal information.

In this sense, the question arises:

1) How can the training of AI systems be made compatible with the principle of necessity, given that this is an activity that often requires the processing of massive amounts of personal data? What safeguards can be adopted to ensure compliance with this principle and enable the proper development of AI systems, while also considering the importance of the quality and diversity of the data used?

While the efficiency and effectiveness of AI systems are closely related to the quality of the oftentimes large amounts of data used during training, systems may vary in the extent to which they rely on personal data. In many circumstances, AI systems will process significant amounts of non-personal data (e.g., agricultural and farming data, environmental data, chemical compounds, geographical/geological data, flight and shipping data, etc.), while some systems may require personal data to perform critical functions, such as reducing the risk of biased outputs.

The development and training of AI systems, and especially those centered on generative AI (genAI), varies in the extent to which it relies on personal data. In some instances, the collection of personal data may be intentional to support critical functions, such as reducing the risk of biased outputs and improving the functionality, security, and quality of the system. In other instances, personal data may be collected incidentally through publicly available sources as part of broader efforts to build rich and diverse datasets. Ultimately, the role that the processing of personal data may play in AI development requires careful contextual and risk-based analysis. While it is important to ensure that the personal data used is accurate, adequate, relevant, and proportionate to the purposes of the training, to ensure proper system functioning and reduce the potential for unintended harms, lawmakers and regulators should avoid overly restrictive and broad requirements and interpretations to exclude personal data from datasets used for AI system development. Many AI systems, especially general-purpose genAI systems, require a considerable amount of data at the development and training stages. In fact, too little data can undermine the development and quality of the system. The questions of how much data is necessary and whether the processing is proportionate during the training and development stages are complex ones that must be considered carefully.

While emerging mitigation measures, including PETs/PPTs, hold promise for lowering the dependence on personal data in the development and training phases, they are still emerging and have challenges. Furthermore, unduly limiting access to data or over-relying on data minimizing methods risks creating negative impacts on the quality of genAI systems and hindering efforts to prevent and mitigate unintended bias.

Data minimization and proportionality concepts in the context of genAI do not mean that only small volumes of data are legitimate in system training. Rather, data minimization in this context should be understood as limiting the amount of personal data used to what is necessary while permitting the appropriate volume of data for the development of a high-quality system and user experience. Stated differently, "data minimization" cannot mean using less data than would be necessary and appropriate to ensure the quality of a genAI system.

2) What good practices and safeguards should be observed in order to define specific purposes and disclose clear and adequate information that is easily accessible to data subjects regarding the processing of personal data carried out during the development and use of AI systems?

Companies that process personal data to develop AI systems can identify a number of safeguards. For example, if a company conducts a data protection impact assessment for AI-related activities, the assessment can help identify potential safeguards. Only after completing a data protection risk assessment can controllers effectively design AI systems that enable the exercise of data subject rights and provide meaningful information to data subjects about the risks associated with the processing of personal data. These steps are essential to ensuring transparency and effectively handling data subject access requests.

Implementing a comprehensive risk assessment system and robust data governance protocols is crucial for data controllers utilizing AI systems to process personal data. These measures help pinpoint potential risks and discern whether the AI's inputs or outputs are personal data. This awareness facilitates crafting AI systems and development processes that efficiently accommodate the management of data subject access requests by ensuring that modifications or deletions can be made without disrupting the AI's operations or discarding existing systems. This may include minimizing the amount of personal data used, precluding specific sensitive data from processing, or intercepting and preventing potential data breaches. Similarly, any outputs identified as risky or derived from sensitive data can be proactively purged. Ongoing monitoring and evaluation of systems are essential to validate that the risk assessments and mitigation strategies remain current and effective.

Furthermore, it is important to note that the processing of personal data for the purpose of developing and training AI systems is distinct from the purpose of operating and deploying such systems. Similarly, when applying data protection principles to AI systems, it is important to recognize the distinct phases of development and deployment: (i) pre-training data collection and pre-processing; (ii) system training (which can include "fine-tuning"); (iii) evaluation; (iv) risk mitigation; (v) deployment; and (vi) monitoring. Additionally, it is important to note that the provision of genAI services often involves processes beyond those mentioned, such as context augmentation and personalization. Lawmakers and regulators should collaborate with developers and deployers of AI systems to clarify the distinctions in duties and responsibilities across these phases, and to distinguish how these phases may vary based on the AI actor's role.

3) How can the principles of purpose and transparency be reconciled with the use of general-purpose AI systems, i.e. systems that can perform a wide variety of different tasks and serve different purposes?

ANPD should recognize the inherently broad purpose of training a general-purpose AI system (GPAI), which is intended to be deployed for a wide range of applications, many of which will be unknown at the time of development. Training is an iterative process, continuing throughout the use of an AI system. Developers may need to collect, retain, and use data beyond the initial training stage. Such ongoing use of data may be necessary to protect against bias, for instance, and to preserve system robustness, accuracy, and security. The use of personal data for the purposes of training GPAI should be recognized as a legitimate and permissible purpose, so long as other accountability measures and safeguards are reasonably and sufficiently implemented. Furthermore, we note that deployment is also a distinct purpose from training and development. Because GPAI system developers may be unable to describe every conceivable use during development, it is incumbent upon deployers to specify the intended purposes of systems they are deploying. To ensure adequate transparency during deployment, deployers should also provide clear explanations of how user-submitted personal data is used to operate their applications and whether user data will be used to train or improve the system. To preserve the societal benefits of GPAI systems, ANPD should interpret purpose specification requirements flexibly during development. It can require developers to provide sufficient transparency measures that indicate the range of applications or tasks that the system is well suited for, given the developer's resources and monitoring abilities. Such documentation should also, when possible and applicable, outline use cases that the developer considers inappropriate or unsuitable for the system, or prohibits as a condition of use.

Transparency measures should allow individuals to understand how their data is being used and enable them to reasonably exercise their data protection rights. However, the ability for organizations to satisfy such requests may be dependent on the context, and the purpose and intended use of the system. The level of detail provided by transparency measures must also be proportionate to the risk posed by the processing, i.e., the greater the risk posed by the processing, the higher the level of transparency that should be offered to individuals. Transparency should not come at the expense of other important factors, such as usability, functionality, and data security, or create additional burdens for users. The level of transparency should be balanced not only with the need to protect IP rights, copyright, and confidential information, but also with the vulnerabilities of systems and the potential net societal benefit that may outweigh individuals' rights. Risk assessments can help organizations properly weigh these considerations.

4) What good practices and safeguards, as well as parameters or criteria, should be considered throughout the life cycle of AI systems to prevent unlawful or abusive discrimination?

Preventing unlawful or abusive discrimination is an ongoing area of research and requires organizations to take a holistic approach. ANPD should ensure that organizations are able to process personal data to the extent necessary to mitigate bias. Regulatory guidance should be drafted or existing legal requirements interpreted to recognize and enable the processing and retention of sensitive personal data for AI system training, as this is necessary to avoid algorithmic bias or discrimination and ensure content safety. In addition, sensitive personal data may be necessary for the training and development of certain AI systems whose sole purpose is based on the processing of sensitive personal data or to deliver benefits to protected categories of individuals (such as accessibility tools, or health systems). Furthermore, as insufficiently diverse or high-quality data sets may lead to biased, inaccurate, or even harmful system outputs, it is imperative that organizations are able to train systems on an extensive range of data and ensure that the system is useful for deployment in a wide range of contexts. It is critical that organizations also implement robust accountability measures and safeguards (e.g., performing risk assessments, ensuring data quality, implementing redress mechanisms, providing appropriate transparency, etc.) to address the potential risks from bias and discrimination.

With respect to genAI in particular, regulatory guidance should be drafted or interpreted so as to enable the responsible processing of sensitive data for bias reduction and content safety, especially where a genAI application may produce a legal or similarly significant effect on an individual. In many cases, this may require the collection of large and diverse datasets and the processing of personal data, including sensitive data, to train accurate and accessible genAI systems that do not unjustly discriminate or perpetrate biases.

To reduce risks associated with the use of sensitive data, organizations should apply privacy-enhancing and privacy-preserving technologies (PETs and PPTs), and filter out, to the extent possible, unnecessary, inadequate, and irrelevant personal and sensitive data before using datasets to train generative AI systems.

BLOCK 2 - LEGAL HYPOTHESES

The LGPD defines legal hypotheses, set out in articles 7 and 11, which authorize the processing of personal data.

Various legal hypotheses can, in different contexts, support the processing of personal data throughout the life cycle of the AI system. Examples include the execution of public policies, the protection of health, the prevention of fraud and the guarantee of the security of the data subject and the execution of contracts.

Among the legal hypotheses that can be used in the context of AI, consent and legitimate interest require further discussion.

Consent presupposes obtaining a free, informed and unequivocal manifestation by which the data subject agrees to the processing of their personal data for a specific purpose, and may be revoked at a later date. Consent can be difficult to apply in practice in some contexts related to AI systems. This is the case, for example, with the collection of publicly accessible personal data through data scraping techniques in order to train AI systems.

In turn, legitimate interest can support the processing of personal data to meet the legitimate interests of the controller or third parties, including the community. To this end, the LGPD requires the adoption of a series of safeguards, including the definition of appropriate transparency measures and the performance of a balancing test, as already addressed by the ANPD in the "Guidance Guide - Legitimate Interest".1 A relevant limitation to the use of the legal hypothesis of legitimate interest in the context of the use of personal data for training AI systems stems from the fact that this legal hypothesis cannot be used to justify the processing of sensitive personal data.

In this sense, the question arises:

5) Can the processing of personal data in the context of AI systems be supported by the legal hypothesis of consent? Under what circumstances? What are the limitations to the use of this legal hypothesis in these contexts and what safeguards should be observed?

As a point of departure, CIPL notes that there are multiple legal bases available for processing under the LGPD and there is no hierarchy among them. Whether the processing of personal data in the context of AI systems can be supported by consent is highly contextual and must be determined on a case-by-case basis. However, there are significant limitations associated with the use of consent. In some cases, depending on the purpose of the AI system, relying on consent can lead to incomplete or non-representative training data which can cause negative downstream effects.

In CIPL's white paper, "The Limitations of Consent as a Legal Basis for Data Processing in the Digital Society", we suggest that consent is not scalable, places a high burden on individuals, and may drive negative impacts on third parties and the beneficial and accountable re-use of data. These concerns are especially acute in the context of AI systems or applications. These limitations not only strain user autonomy but also have the potential of diminishing the efficiency and effectiveness of the digital ecosystem. Furthermore, it may not be technically feasible or overly burdensome to satisfy individuals' withdrawal of consent or erasure requests. This is because tracing the data back to specific individuals is not only resource-intensive and nearly impossible with third-party data, but also risks individuals' privacy and may even structurally compromise AI systems, significantly impacting the functioning of AI system.

To foster innovation while protecting individuals, CIPL recommends that the legitimate interest method, combined with robust organizational accountability measures and data protection impact assessments, offers a more flexible and balanced approach in the context of AI systems. Significantly, this empowers organizations to process data for purposes beyond the initial point

of collection while ensuring individual protection through transparency, the right to object, and a contextual risk-based approach.

6) Can the processing of personal data in the context of AI systems be supported by the legal hypothesis of legitimate interest? Under what circumstances? If so, what safeguards should be adopted in these situations with a view to protecting the rights of data subjects, especially considering the prohibition on processing sensitive personal data based on the legal hypothesis of legitimate interest? In particular, can the collection of personal data for the training of AI systems, especially through data scraping techniques, be based on the legal hypothesis of legitimate interest?

Legitimate interest is available as a potential legal basis for scraping and processing publicly available personal data. Using the legitimate interest basis for training AI systems balances the societal benefits and public interest of innovation and AI with the appropriate guardrails to protect individual rights. Importantly, the legitimate interests concerned could be the data controller's, users', or society's at large. The LGPD, like the EU GDPR and UK GDPR, requires organizations to complete a balancing test to determine whether the legitimate interest lawful basis applies to their intended processing.

The recent Opinion 28/2024 published by the EDPB described conditions for the application of legitimate interest as a legal basis for the processing of personal data for the development and deployment of AI systems. Furthermore, the Court of Justice of the European Union (CJEU) recently affirmed an organization's "commercial interest" as a legitimate interest so long as the commercial interest is lawful and complies with the EU GDPR's three-part balancing test. This judgment is significant in that it presents a framework for responsible innovation through the legitimate interests balancing test; stakeholders should carefully review its applicability to the development of AI systems, and genAI, in particular. The CJEU has also ruled that broader, socioeconomic interests rather than the controller's own interest may satisfy the criteria necessary to prove a legitimate interest. Similarly, the UK ICO has long supported a broad interpretation of what can constitute a legitimate interest, including societal benefits. It is worth noting that LGPD Paragraph 4 of Art.7 provides for the exemption from the requirement of consent for data made manifestly public by the data subject. Publicly available data is at the core of how many genAI systems are trained and is foundational to both system quality and functionality. The French CNIL[5], the EDPB[6], and the UK ICO[7] have suggested that web scraping

[5] CNIL, "The legal basis of legitimate interests: Focus sheet on measures to implement in case of data collection by web scraping", July 2, 2024, https://www.cnil.fr/en/legal-basis-legitimate-interests-focus-sheet-measures-implement-case-data-collection-web-scraping

[6] EDPB, "Report of the work undertaken by the ChatGPT Taskforce", May 23, 2024, https://www.edpb.europa.eu/system/files/2024-05/edpb_20240523_report_chatgpt_taskforce_en.pdf

[7] UK ICO, "The lawful basis for web scraping to train generative AI models", https://ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/response-to-the-consultation-series-on-generative-ai/the-lawful-basis-for-web-scraping-to-train-generative-ai-models/

on the basis of legitimate interests might be possible. ANPD should also recognize fairness and reduction of bias and discrimination as legitimate interests.

Relying on legitimate interest requires organizations to consider several factors, including:

- industry practices, such as the robots.txt protocol (which provides directives on what parts of the website can system
- be scraped);
- website policies and technical measures that prohibit web scraping;
- intellectual property and contract laws;
- whether data is made public, as discussed above;
- filtering out, to the extent possible, unnecessary, inadequate, and irrelevant personal and sensitive data before using datasets to train genAI systems; and
- where up-front filtering is impractical, data annotation to mark out identifiable personal and sensitive data inputted in AI training, with technical measures to withhold it in system outputs in case of attempts to draw out or 'regurgitate' it.

BLOCK 3 - RIGHTS OF DATA SUBJECTS

The use of personal data in AI systems can have significant impacts on the rights of data subjects.

One of the most critical aspects concerns the making of decisions based solely on the automated processing of personal data and which can produce legal effects or significantly impact the interests of individuals. The LGPD establishes that the data subject has the right to understand the criteria used for this decision and, more specifically, to request a review when such decisions affect their interests. This seeks to avoid or mitigate possible errors, biases or unlawful or abusive discrimination that may arise from automated decisions that negatively affect the individual.

The use of personal data in AI systems also requires careful consideration of the possible negative impacts on the data subject, which may occur in decisions aimed at defining their personal, professional, consumer and credit profile or aspects of their personality.

Still in relation to the exercise of rights, we highlight the confirmation of the existence of processing, access to personal data, the correction of incomplete, inaccurate or outdated data, the anonymization, blocking or deletion of unnecessary, excessive data or data processed in disagreement with the provisions of the law, as well as the possibility of revoking consent. The data subject can also object to processing carried out on the basis of one of the hypotheses for waiving consent, in the event of non-compliance with the provisions of the LGPD.

Compliance with the LGPD not only establishes a framework of protection for data subjects, but also strengthens their confidence in the development and use of AI systems, ensuring that technological progress is always aligned with the protection of fundamental rights and privacy.

In this sense, the question arises:

7) How do the rights of the data subject, provided for in the LGPD, apply to AI systems?

The LGPD applies to personal data processing activities, regardless of the means by which they are carried out. In this sense, the LGPD does not apply to AI systems per se, but rather to the processing of personal data carried out by them. It is important that the ANPD avoid creating overly specific requirements on the methods by which a data subject exercises these rights, given the wide variety of controllers that must comply with rights requests across a range of products and services. The best methods of communication to exercise these rights vary greatly across different products and services, including those related to AI.

Where personal data is collected directly from individuals, the organization collecting it must explain at the time of collection how the data will be used and how individuals can exercise their data rights. Instead of individual disclosures, organizations should be able to fulfill transparency and notice requirements through public disclosures and information campaigns, accessible privacy notices, or other informational resources explaining how data is used in the context of the system. The responsibility to inform individuals about the use of their data should fall to the entity closest to the individual from whom the data is collected, whether that be during development or deployment. For example, in the context of genAI, a deployer client of the system developer who provides personal data to the developer for training purposes is closer to the individual than the developer. Deployers should also be responsible for complying with access requests received concerning the personal data they process during their particular deployment of a genAI system.

Erasure: Lawmakers and regulators should consider that, in the case of web-scraped data used during training but not cataloged or further filtered to identify personal data, it may be unreasonable or technically infeasible for a developer to respond to requests for erasure, particularly after the data has been used for training. Furthermore, there may be instances where organizations are unable to comply with erasure requests because the associated data is subject to data retention requirements from other legal acts, such as anti-money laundering requirements, or is under hold due to litigation proceedings and is thus, prohibited from being further processed, including for deletion or modification of the data. ANPD should also allow organizations to process personal data to the extent necessary to mitigate bias or meet other legal requirements, such as those related to security and transparency.

Objection: Organizations that rely on the legitimate interest legal basis should allow individuals to object to the use of their personal data for system operation, development, and improvement in an accessible manner, and cease processing it for future system versions unless the organization can demonstrate compelling interests that override the reasons for the objection.

8) What are the good practices and safeguards to be observed when providing service channels for data subjects to exercise their rights, such as the rights of access, opposition and review of automated

decisions, in the context of the processing of personal data by AI systems? If possible, describe the tools used to implement such service channels, with the respective parameters used.

The LGPD does not require creation of service channels specific to each AI system or to AI generally. In 2024, CIPL published a report, "Building Accountable AI Programs: Mapping Best Practices to the CIPL Accountability Framework", that noted that some organizations were beginning to explore creation of such dedicated channels for AI services, others were using established communication channels to enable individuals to file complaints, report issues, or ask questions regarding their rights as users of an AI service.

Principles that our members prioritize for such service channels include ensuring that service channels are easy to access; are indicated in terms of service; include clear deadlines for how to use; have service logs to ensure traceability and accountability; prioritize timely response consistent with legal requirements; and are staffed by employees trained in how to deal well with requests.

9) Should there be specific safeguards and limits for the processing of sensitive personal data and for the processing of personal data of children, adolescents and the elderly during the life cycle stages of AI systems?

Data protection laws, including the LGPD, generally place stricter rules on sensitive data processing, such as through specific consent requirements. Such requirements can place organizations in a position of having to exclude sensitive data from training datasets to the detriment of the performance of an AI system, including with respect to groups whose data is afforded special protections under the LGPD, where such consent is not obtainable for example. For example, not being able to rely on the use of sensitive data may hamper efforts to reduce bias, improve system fairness and quality, or promote content safety, which will often rely on the processing of sensitive data. Thus, laws and regulatory guidance should be drafted or interpreted so as to enable the responsible processing of sensitive data for bias reduction and the promotion of content safety. This should also be the case where an AI application may produce a legal or similarly significant effect on an individual.

At the same time, limiting the processing of personal and sensitive data in instances where it is unnecessary for system performance can help mitigate risks associated with AI systems. Employing privacy-enhancing and privacy-preserving technologies (PETs/PPTs) such as anonymization, synthetic data, data annotation techniques, and differential privacy may in some cases be able to provide AI systems with sufficiently diverse data during training while reducing the risks associated with the use of sensitive data. For instance, prior to the development stage, safeguards in compliance with privacy-by-design principles may be considered, such as filters or pattern recognition algorithms, to reduce the amount of personal data in any downstream output; synthetic data may be used in some instances to train or validate the system without exposing sensitive information; differential privacy may, in certain circumstances, be used to add noise during training to prevent identification; and homomorphic encryption can keep data

secure throughout the training process. AI systems can also be valuable tools in scrubbing datasets of personal and sensitive data prior to training. The effectiveness and usefulness of these mitigation measures is subject to a case-by-case analysis and determination. Additionally, these mitigation measures are subject to rapid technological development and organizations should be provided with sufficient flexibility to apply mitigation measures when doing so is demonstrably beneficial.

10) What requirements must be met in order to guarantee and enforce the right to review automated decisions (Art. 20 of the LGPD)? What can be considered a decision taken solely on the basis of automated processing of personal data? What interests could be affected?

Article 20 of the LGPD provides for a right to request a review when decisions that are taken "solely on the basis of automated processing of personal data" affect the data subject's interests. CIPL suggests that a data subject's interests are affected when the automated decision produces legal or similarly significant effect. This risk-based interpretation allows organizations to focus more resources on automated decision-making systems without the additional burden of expending heightened resources on low-risk automated decision-making.

Examples of automated decisions producing legal effects could include decisions affecting the legal status of individuals; decisions affecting legal entitlements of individuals; decisions affecting legal and public rights of individuals; decisions affecting contractual rights of individuals; and decisions affecting individuals' rights of ownership. Depending on applicable laws, decisions producing similarly significant effects to those of legal decisions include decisions affecting an individual's eligibility and access to essential services — e.g., health, education, banking, insurance; decisions affecting school and university admissions; and decisions affecting an individual's promotion or pay. Low-risk automated decisions that typically do not produce legal or similarly significant effects include decisions ensuring network, information and asset security and preventing cyber-attacks; decisions related to fraud detection and prevention; decisions related to commonly accepted forms of targeted advertising; and decisions to automatically disconnect a service when customers fail to make timely payments.

11) In what cases and under what conditions might human review of automated decisions be necessary in order to adequately guarantee the rights of data subjects?

Human review may be useful when automated decisions produce legal or similarly significant effects. At the same time, human review is not foolproof, and circumstances may warrant consideration of additional automated systems as part of review processes. Where humans are involved in the review of automated decisions, they should be equipped with the proper training, knowledge, and experience to capably review automated decisions.

12) What are the parameters to be observed for the provision of clear and adequate information regarding the criteria and procedures used for the automated decision, under the terms of §1 of art. 20 of the LGPD? What limits and parameters of commercial and industrial secrecy justify not complying with the provision of information, as set out in the same legal provision?

Regulations should provide organizations with sufficient flexibility to respond to requests for review. Responses should be meaningful and proportionate to the level of risk associated with the automated decision and flexible enough as to not require organizations to disclose trade secrets or intellectual property. Regulatory guidance should assist organizations with finding ways to provide individuals with simple and clear information about the rationale behind an automated decision without requiring a complex explanation of the algorithms used or the disclosure of the full algorithm.

Providing meaningful automated decision-making transparency is contextual and regulations should be flexible enough to accommodate different contexts without requiring businesses to disclose trade secrets, intellectual property, or internal security and safety mechanisms (such as anti-fraud and cybersecurity tools). Regulations should make clear that organizations are not expected to disclose trade secrets, intellectual property, or internal security and safety mechanisms.

As detailed in CIPL's report, "Building Accountable AI Programs: Mapping Best Practices to the CIPL Accountability Framework", accountable organizations are implementing a wide range of practices to ensure clear and adequate information is provided to individuals that request a review of automated decisions. These practices include establishing AI ethics oversight bodies or committees to review automated decision-making systems that produce legal or similarly significant effects and produce guidelines for clear and adequate information sharing when individuals exercise their right to review such decisions.

BLOCK 4 - GOOD PRACTICES AND GOVERNANCE

Governance and the adoption of good practices in the use of personal data in AI systems can be a good strategy for processing agents to ensure the protection of data subjects' rights and compliance with the LGPD.

The adoption of appropriate security measures that are compatible with the risk involved in each situation, in order to avoid the occurrence of security incidents and mitigate possible negative impacts on data subjects are mechanisms provided for in the LGPD system that must be adopted throughout the life cycle of AI systems that use personal data. In this sense, art. 46, § 2, of the LGPD states that security measures must be observed from the product or service design phase to its execution, a rule that is also applicable to the context of the development and use of AI systems.

Similarly, art. 50 of the LGPD states that processing agents may formulate rules of good practice and governance that establish the conditions of organization, the operating regime, procedures, including

complaints and petitions from data subjects, security standards, technical standards, specific obligations for the various parties involved in the processing, educational actions, internal supervision and risk mitigation mechanisms and other aspects related to the processing of personal data. In addition, §2, I of the same article establishes the minimum requirements for the implementation of the privacy governance program.

In this respect, it is important to note that the use of safeguards, such as anonymization techniques, can provide greater protection for data subjects, allowing information to be processed without it being associated with a specific individual. Anonymization is a technique that helps to minimize risks and protect privacy, especially in the training of AI systems, where large volumes of data are used and may contain information that allows for potential risks to the civil liberties and fundamental rights of data subjects.

Another relevant mechanism provided for in the LGPD that can be used for governance and risk management in the context of the development and use of AI systems is the personal data protection impact report (DPIA). If well prepared, an RIPD can provide the organization with a suitable tool for understanding the risks involved and the mitigation measures adopted in the case, as well as enabling accountability for the system.

In this sense, AI governance involves, among other actions, the implementation of policies and processes that ensure comprehensive compliance with standards and good practices regarding the protection of personal data and guide how personal data should be collected, processed, stored and used throughout the life cycle of an AI system.

The question therefore arises:

13) How can privacy governance programs be used as a mechanism to promote compliance of the development and use of AI systems with the LGPD? What requirements, specifically relating to the development and use of AI systems, should be observed in such cases?

> In CIPL's report on "Building Accountable AI Programs: Mapping Best Practices to the CIPL Accountability Framework", we suggest that AI governance works best when it leverages knowledge from other disciplines within an organization, including data protection, information security, human rights and ethics, and more. Many organizations may have existing privacy/ data protection governance programs that can and should be leveraged to develop responsible and accountable AI governance programs.
>
> Given that privacy is a common thread connecting multiple data-rich business functions (e.g., security, HR, marketing, product, emerging technologies, business innovation, policy, etc.), Chief Privacy Officers (CPOs) and their teams have become a natural first point of contact for various AI compliance and regulatory queries. Many executives believe that data privacy teams are uniquely positioned to shoulder these new AI responsibilities and lead the charge in growing and maturing organizations' AI accountability programs. First, with ongoing AI regulatory developments and numerous others forecasted soon, data privacy professionals are already accustomed to implementing evolving laws that impact data and technology use and

considering multiple, overlapping regulatory schemes. Second, many accountable AI programs draw from practices familiar to data privacy management programs, such as data protection impact assessments and risk assessments. Moreover, the framework for implementing data privacy programs—such as the elements of accountability and many of the controls, tools, and processes— can be replicated in responsible AI programs. Data privacy professionals can thus leverage their experience to create consistent structures for accountable governance of AI.

In recent years, many CPOs have started to expand their roles to include a broader remit of data and digital ethics, data regulation, and digital trust, which many believe should naturally include AI. Organizations are assessing how emerging AI governance programs can leverage and build upon existing data protection and privacy processes instead of creating new frameworks from scratch.

14) Considering the principle of responsibility and accountability, what information should be documented during the life cycle of an AI system? In which specific contexts related to AI systems is it recommended to draw up a DPIA? In this case, is it possible to establish specific requirements to be observed when drawing up the DPIA

Insofar as AI system development and deployment involve the processing of personal data, applicable DPIA requirements under the LGPD apply. The risk associated with the processing of data is highly contextual, and organizations should be required to conduct rudimentary risk assessments for all processing activities, even presumptively low-risk processing. Such initial, rudimentary risk assessments, coupled with guidance on what might be high-risk activities, could trigger more robust, full-blown data privacy impact assessments where a likelihood of a higher risk is identified or expected.

Regulators can provide organizations with a risk taxonomy that reflects the requirements of the LGPD, which could enable organizations to implement a risk-based approach and establish a methodology when assessing risks to individuals. Any risk classification presumptions should be rebuttable through documented impact assessments that regulators could review.

Many organizations will want to streamline data protection, privacy, and AI impact assessments. Therefore, the ANPD should maintain a flexible approach so long as all substantive elements are included based on the context of the underlying processing activities. Further, to promote interoperability between jurisdictions, the ANPD should encourage the development of documentation in the context of AI systems as a best practice, based on international best practices and standards, and even cooperate with other data protection authorities to create data protection impact assessment templates to guide businesses when they must bridge legal and technical differences between legal systems.

15) Considering the life cycle of an AI system, at what point and in what context would anonymization be feasible or necessary? What technique would be used? What other security measures could possibly be used to protect the privacy of data subjects?

As a forthcoming CIPL white paper documents, privacy-enhancing and privacy-preserving technologies (PETs and PPTs) can play a critical role in safeguarding individual privacy during the development of AI systems. These include synthetic data, homomorphic encryption, differential privacy, federated learning and analysis, and secure multi-party computation.[8] Regulators and data-using organizations across many jurisdictions are exploring the extent to which PETs can satisfy legal standards for anonymization of data. In Europe, a long-held view of the regulators was that encryption does not meet the standard of anonymization.[9] However, in a recent case, the European General Court ruled that in order to determine whether an individual is identifiable, account should be taken of all means reasonably likely to be used, and that this test must be performed from the perspective of the recipient/holder of the data.[10] This ruling has been interpreted by some to mean that if the decryption key is inaccessible, then the data could be deemed anonymous. Regulatory clarity on this point and on the ability of PETs to anonymize data could be a powerful incentive for using PETs more broadly. Regulators should take a risk-based approach to what the legal threshold for anonymization is, viewing it not as a reduction of the risk of re-identification to zero, but rather to a sufficiently low level, taking into account the context and purpose of the data processing.

[8] CIPL, *PETs and PPTs in AI: Operationalizing Privacy by Design and Default* (forthcoming).
[9] Opinion 05/2014 on Anonymisation Techniques, Article 29 Data Protection Working Party, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.
[10] *Single Resolution Board v. European Data Protection Supervisor* (Case T-557/20), https://eur-lex.europa.eu/legal-content/EN/TXTPDF/?uri=CELEX:6202T%3A2023%3A219.