Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Applying Data Protection Principles to Generative AI: Practical Approaches for Organizations and Regulators

Discussion Paper

December 2024

Summary of Recommendations

*The following CIPL **recommendations** are intended to facilitate the application of data protection principles to generative AI (genAI) models and systems that process personal data. They are discussed in greater detail in Section II of this paper.*

1. To enable beneficial development and use of AI technologies in the modern information age, laws and regulatory guidance should **facilitate lawful mechanisms for the use of personal data in model training**. Lawmakers and regulators should **avoid legal interpretations that are unduly restrictive** regarding the use of personal data in AI model training, development and deployment.

2. Different data privacy rules, considerations, and mitigations apply in different phases of the AI lifecycle— data collection, model training, fine tuning, and deployment. **Regulators and organizations should interpret data protection principles separately in the context of each relevant phase of the AI technologies**.

3. Organizations should be able to rely on the **"legitimate interests" legal basis** for processing publicly available personal data collected through **web scraping and** personal data that they **already have in their possession and control (first-party data)** for genAI model training, as long as the interest concerned (which could be the controller's, users', or society's at large) is not outweighed by the fundamental rights of individuals and appropriate, risk-based mitigation measures are put in place.

4. Laws and regulatory guidance should be drafted or interpreted to recognize and enable the **processing and retention of sensitive personal data for AI model 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).

5. Developers should explore opportunities to employ **privacy-enhancing and privacy-preserving technologies (PETs/PPTs),** such as synthetic data and differential privacy. This would enable genAI models to have the rich datasets they need during training while reducing the risks associated with the use of personal data. Laws and regulatory guidance should **encourage the use of and acknowledge the need for continued research and investment in PETs/PPTs**.

6. The **fairness principle** is useful in the genAI context and should be interpreted to facilitate personal data processing in genAI model development to train **accurate and accessible models that do not unjustly discriminate**. Consideration of fairness also need to take into account the impact on the individual or society of not developing a particular AI application.

7. **Data minimization** should be understood contextually, as **limiting the collection and use of data that is necessary for the intended purpose** (e.g., model training, model fine-tuning, or model deployment for a particular purpose). Data minimization should not stand in the way of enabling the collection and use of data that is necessary and appropriate for achieving a robust and high-quality genAI model. As such, this principle does not prohibit or conflict with the collection and use of large volumes of data.

8. **Training general-purpose AI models should be recognized as a legitimate and permissible purpose in itself**, so long as appropriate accountability measures and safeguards are reasonably and sufficiently implemented.

9. **Purpose or use limitation principles should be sufficiently flexible**: In the context of genAI, purpose limitation principles in laws and regulations should allow organizations to articulate data processing purposes that are sufficiently flexible for the range of potentially useful applications for which genAI models may be used. Furthermore, processing personal data for the development of a genAI *model* should be treated as a **separate purpose** from processing personal data for the development, deployment or improvement of a specific *application that uses a genAI model*.

10. 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. Where data is not collected directly from individuals, organizations should be able to fulfil transparency requirements through **public disclosures** or other informational resources.

11. Where appropriate and practicable, individuals should be able to **request that their input prompts and model output responses not be included in genAI model fine-tuning**, especially if such prompts include personal or sensitive data.

12. **Transparency in the context of genAI models should be contextually appropriate and meaningful**, while also fulfilling transparency requirements under applicable laws and regulations. Transparency should not come at the expense of other important factors, such as usability, functionality, and security, or create additional burdens for users. Organizations should also consider transparency in the wider sense, beyond individuals and users—to regulators, auditors, and red-team experts.

13. Lawmakers and regulators should consult with developers and deployers of genAI systems to **clarify the distinctions in duties and responsibilities** across the phases of genAI development.

14. Organizations developing and deploying genAI models and systems must invest in **comprehensive and risk-based AI and data privacy programs**, continually improving and evolving their controls and best practices. Lawmakers and regulators should encourage and reward organizational accountability in development and deployment of AI, including the existence and demonstration of AI and data privacy management programs.

Table of Contents

I. Introduction and Background

Generative AI (genAI) systems have arrived and are here to stay, supporting individuals and enterprise users in generating audio, code, image, text, and video content at scale and with speed. In the short time that genAI tools have been available for broad public use, we have witnessed widespread adoption by individuals and organizations around the world. OpenAI's ChatGPT now has more than 200 million weekly active users,[1] Microsoft's Github Copilot has over one million paying users,[2] and, according to a 2024 study by the McKinsey Technology Council, 65% of global organizations have adopted genAI systems in at least one business function.[3]

As a general rule, genAI systems rely on general-purpose artificial intelligence (AI) models also called foundation models[4] that are usually trained on vast amounts of data to achieve a variety of purposes. For example, large language models are trained on billions of bytes of text data from a multitude of sources such as publicly available data from the web (which can include personal data), licensed data, and academic and industry datasets.[5] From these large and diverse datasets, genAI models are trained to recognize statistical relationships between words and other data, such as images, videos, and audio, in response to a wide range of user prompts and make probabilistic predictions which generate useful outputs.[6] In addition , genAI models can be further "fine-tuned" and personalized with specifically curated data to perform better for a specific identified purpose. For example, a genAI model can be fine-tuned with medical data to assist physicians and healthcare workers with note-taking and clinical documentation.[7] A model can also be personalized to answer novel questions in a customer engagement or individualized tutoring context.

GenAI systems require users to input prompts to obtain a generated output, and inputs and outputs may sometimes include personal or even sensitive information.[8] During deployment, genAI models can leak or disclose personal data from training datasets and generate inaccurate data related to individuals (also known as a "hallucination"), and malicious actors can use a variety of methods to bypass protective guardrails set up to avoid disclosure of personal data from genAI models. As a result, data protection authorities, other regulators, as well as researchers, are increasingly discussing whether and how[9] data protection laws apply to genAI tools, what new risks for data protection may arise from these systems, and how potential tensions between certain data protection principles and genAI might be resolved.[10]

This discussion paper considers the following key privacy and data protection concepts and explores how they can be effectively applied to the development and deployment of genAI models and systems:[11]

- fairness;
- collection limitation;
- purpose specification;
- use limitation;[12]
- individual rights;
- transparency;

- organizational accountability; and
- cross-border data transfers.

The recommendations do not apply to non-personal data used to develop, train, fine-tune or deploy genAI models and systems.

The analysis in this paper builds on the Centre for Information Policy Leadership (CIPL)'s[1] previous work on the intersection of data protection, artificial intelligence, and organizational accountability and synthesizes it for the context of genAI models and systems.[13] Our work has included convening global regulators, academia, and industry to discuss the emerging tensions between AI technologies and data protection principles and develop potential solutions that resolve or mitigate these tensions. It has also resulted in numerous CIPL reports, white papers, and public consultation responses.[14]

II. Discussion

A. Distinguishing the phases of genAI development

When applying data protection principles to genAI, it is important to recognize the distinct phases of genAI development and deployment: (i) pre-training data collection and pre-processing; (ii) model 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.

B. The role of personal data in genAI model development

GenAI development 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 model. 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 genAI development requires careful contextual and risk-based analysis. To ensure proper model functioning and reduce the potential for unintended harms, lawmakers and regulators should therefore

[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.

avoid overly restrictive and broad requirements to exclude personal data from datasets used for genAI model development.

C. The status of "sensitive data" in genAI model development

The role that the potential processing of sensitive data, or special category personal data,[15] plays in genAI model training is the subject of increased scrutiny and attention. Data protection laws generally place stricter rules on sensitive data processing, such as through specific consent requirements. This can place organizations in a position to potentially exclude sensitive data from training datasets to the detriment of the performance of the model, where such consent is not obtainable for example.[16] This may unintentionally hamper approaches to reduce bias and improve model fairness or content safety, which will often rely on the processing of sensitive data.[17] At a minimum, this dilemma suggests that laws and 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.

D. Mitigations, PETs and PPTs

At the same time, limiting the processing of personal and sensitive data in instances where it is unnecessary to model performance can help mitigate risks associated with genAI models. Employing privacy-enhancing and privacy-preserving technologies (PETs/PPTs) such as anonymization, synthetic data, and differential privacy may in some cases be able to provide genAI models with sufficiently diverse data during training while reducing the risks associated with the use of sensitive data.[18] 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 model 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. GenAI systems can also be valuable tools in scrubbing datasets of personal data prior to training.

E. Specific data protection principles

a. Fair Processing—*This principle requires organizations to ensure that personal data is processed fairly. It means that organizations are required to consider the impact of the processing and avoid unfair consequences or outcomes from that processing. Fair processing is often linked to transparency, ensuring data processing is not deceptive or otherwise surprising to the individual. To comply with the fairness principle, genAI models and systems cannot produce unfair, discriminatory, or biased outcomes. There may be tensions surrounding whether the processing of personal data in genAI model development is conducted fairly.*

Fairness and fair processing are common principles in global data protections laws.[19] Despite the fairness principle's importance, it is often not authoritatively or consistently defined and is analyzed in a case-by-case basis. In practice, fairness can appear as an amorphous concept that is subjective, contextual, and influenced by a variety of social, cultural, and legal factors.

The difficulty and importance of defining and ensuring fairness are only magnified in the genAI contexts. At the same time, fairness requirements have increasingly surfaced in laws and regulations targeting automated systems (which may rely on genAI models). Laws and regulations should therefore facilitate and encourage organizations' ability to process personal data to meet fairness requirements. 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 models that do not unjustly discriminate or perpetrate biases.

Laws and regulations should also acknowledge that the fairness principle requires organizations developing and deploying genAI to balance the various rights, freedoms, and interests involved with the development and use of the technology. For example, organizations developing genAI should be encouraged to take into account the potential societal benefits of a genAI model developed on representative data when weighing the risks of processing an individual's data, provided that proper mitigations are put in place.

b. Collection Limitation Principle/ Legal Basis for Processing—*This core data protection principle limits the collection of personal data to what is necessary for a specific purpose (like the concepts of **data minimization** and **proportionality**) and prescribes lawful means to do so. In many jurisdictions the collection of personal data may require a **legal basis** such as consent, contractual necessity, public interest, or the legitimate interests of the controller or a third party. In jurisdictions that do not specify legal bases, data collection must frequently still follow other requirements, such as purpose specification and fairness, to be lawful.*

Both the novelty of genAI models and the collection of large amounts of existing data required to train genAI models are not necessarily at odds with existing interpretations of collection limitation principles. Furthermore, existing legal frameworks governing data collection may be ill-suited for the training of genAI models if interpreted too narrowly.

Collecting data through web scraping: Many genAI models are developed to achieve a broad range of tasks, and the scraping of publicly available data from the web is a common practice for developers to gather large and diverse datasets. These large datasets are often necessary to train genAI models on patterns and relationships between data to generate reliable and fair results. Web-scraped data can include personal and even sensitive data, which requires the consideration of privacy and data protection laws. Responsible web scraping requires appropriate guardrails to ensure data minimization and accountable processing. However, this should not be a one-sided effort, and websites themselves should implement technological solutions and appropriate terms to prevent or limit scraping where necessary.[20]

As significant volumes of data are often required to effectively train genAI models, publicly available personal data on the Internet is often critical to the functioning of many AI models. For example, a user may ask "who is the Prime Minister of Spain" and it would be illogical for the AI model to respond that the information is redacted to protect privacy. The ability to web scrape publicly available data is also important to allow smaller players who have less access to first party data to develop AI models. However, web scraping of personal data should go hand in hand with proper limitations and precautions, and be proportionate to the processing goals.

Data protection laws that include legal bases for processing typically include a "legitimate interests" legal basis, which is often particularly relevant and appropriate in the context of generative AI model development. Using the legitimate interests basis for training genAI models balances the societal benefits and public interest of innovation and genAI use with the appropriate guardrails to protect individual rights. Importantly, the legitimate interests concerned could be the controller's, users' or society's at large. For example, under the GDPR, including the UK GDPR, organizations must complete a three-part balancing test to determine whether the legitimate interests lawful basis applies to their intended processing:

1) is there a specific, clear, and valid commercial or societal interest for processing;
2) is the processing necessary to achieve the identified interest; and
3) do the interests and rights of individuals override the interests of the organization.[21]

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 that three-part balancing test.[22] 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 genAI. The CJEU has also ruled that broader, socio-economic interests rather than the controller's own interest may satisfy the criteria necessary to prove a legitimate interest.[23]

Certainly, there are broader societal benefits derived from training genAI models on a large variety of personal data, such as easing information and technology access, enhancing genAI's diversity and accessibility (e.g., reflecting local languages, cultures, etc.), and avoiding the exclusion of certain groups from genAI innovation.

In cases where web scraping results in the collection of sensitive data, regulators and lawmakers should ensure that there is a lawful basis to collect and process this information for legitimate purposes, such as ensuring accurate and non-biased results and content safety. The European Union Artificial Intelligence Act rightly acknowledges this in Article 10(5) by providing certain AI developers with a lawful means to process sensitive data to "ensur[e] bias detection and correction" provided certain conditions are met.[24] Lawmakers and regulators in other jurisdictions that heavily regulate the collection of sensitive data should consider amending their privacy and data protection laws to allow for similar processing. At the same time, model developers and deployers should implement appropriate controls to prevent the disclosure of such data types.

In other cases, genAI developers may only incidentally collect personal or sensitive data as a result of web scraping. In these cases, incidental collection should not be inherently unlawful where organizations take measures to filter out unnecessary data and apply technical and organizational protections to mitigate against risks. Requiring organizations to identify all personal or indeed sensitive data in pretraining datasets, given the scale and structure of the datasets involved, would require substantively more processing and create more risk. Rather, organizations should be expected to take reasonable measures, especially at the output stage, to protect the privacy

rights of individuals and prevent harms arising from processing of personal data.[25] The Court of Justice of the European Union (CJEU) has previously addressed challenges in adapting data protection principles to novel technologies. In the *Google Spain* ruling, the Court handled the lawful use of web-sourced data. Similarly, in *GC and Others v. CNIL*, the Court addressed the incidental collection of sensitive data in the context of search engine and web data.

Additionally, organizations should be able to web scrape sensitive data that individuals chose to make public. An important parallel can be drawn from the case of *Meta Platforms Inc. v. Bundeskartellamt* where the CJEU considered the legitimate interests of a controller to process sensitive data and when such data could be considered publicly accessible.[26] The CJEU clarified that personal data is manifestly made public in cases where an individual "intended, explicitly and by a clear affirmative action, to make the personal data in question accessible to the general public."[27] Where an individual has available to them settings and is provided with full knowledge that their information can be accessed by the general public or limited to a select few and explicitly selects to make their information public, then the individual has manifestly made their information public.[28] As a general matter, genAI model developers should be able to use such data so long as the developer's legitimate interest is not outweighed by the rights of individuals.

Other exceptions that allow the processing of sensitive personal data under the GDPR include, in particular, processing that is "necessary for reasons of substantial public interest, on the basis of Union or Member State law"[29] and processing that is necessary for scientific or historical research purposes.[30]

In jurisdictions that recognize the "legitimate interests" legal basis, lawmakers and regulators can recognize fairness and the reduction of bias and discrimination as legitimate interests. Regulators and lawmakers should consider how the use of PETs on web scraped datasets can reduce risks associated with the collection or processing of personal data while preserving beneficial outcomes such as bias reduction. Relying on legitimate interests as a legal basis also requires organizations to implement demonstrable policies and procedures to mitigate the potential risks and harms to individual rights.

Relying on legitimate interests to web scrape for training data requires a case-by-case analysis and organizations should consider several factors, including:

- industry practices, such as the robots.txt protocol[31] (which provides directives for web scrapers on what parts of the website can and cannot be scraped);
- website policies and technical measures that prohibit web scraping (i.e., some websites permit web scraping while others have terms of use that prohibit web scraping);
- intellectual property and contract laws;
- whether data is made public (e.g., web scraping tools should avoid websites that are password protected, require log-in credentials, or are behind paywalls);
- applying PETs and PPTs when possible, to mitigate risks associated with personal and sensitive data; and

- filtering out, to the extent possible, unnecessary, inadequate, and irrelevant personal and sensitive data before using datasets to train genAI models, in compliance with data minimization principles.

Processing first-party personal data for genAI development: Laws and regulations should equally enable organizations to process first-party personal data to develop genAI models, provided that proper transparency and risk mitigation measures are in place. As in the web-scraping context, the legitimate interest legal basis should also be available in the first-party context in jurisdictions that recognize this legal basis. However, additional accountability measures may be warranted for some data (e.g., sensitive data)—such as increased transparency and opt-in or opt-out controls—when organizations are developing genAI models.

Unduly impeding an organization's ability to process first-party personal data for legitimate interests may hinder responsible genAI model development. For example, in some cases an organization may only be able to provide a genAI model in a certain language if it is able to process first-party data because third-party sources do not provide sufficient or accurate data in that language. It will be important in these cases for organizations that use first-party data to train genAI models to provide users with easily accessible and proper notice about how their data will be used in addition to robust mitigations such as opt-out mechanisms where appropriate. Other mitigations could include measures such as excluding or filtering out the first-party personal data of children.

Data minimization and proportionality: As noted, data minimization and proportionality concepts go to the same issue as "collection limitation"—reducing and limiting the processing of personal data to what is necessary and proportionate for the purpose pursued. Many genAI models, especially general-purpose genAI models, 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 model.

The questions of how much data is necessary and whether the processing is proportionate during the training and development stages and how best to implement data minimization measures in these contexts, including the role of PETs/PPTs, are complex ones that must be considered carefully. While emerging mitigation measures, such as synthetic data, hold promise for lowering the dependence on personal data in the development and training phases, they are still emerging and have challenges.[32] Furthermore, unduly limiting access to data or over-relying on data minimizing methods risks creating negative impacts on the quality of genAI models 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 model 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 model 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 model.

With this in mind, it is possible in some cases for controllers to redact personal data that could be used to link data to an individual and leverage synthetic data in its place. Controllers can also limit personal data processing to the stage for which it is necessary (e.g., during fine-tuning stages as opposed to general model training) and excluding from datasets the personal data of authenticated users under a certain age. These measures must always be balanced with other legal obligations, including the provision of safe and reliable products.

It is also important to note that data minimization and proportionality should not just be interpreted in the context of training a genAI model, but should also be considered in other genAI processes. For example, implementing appropriate controls at the output stage, such as output filters (discussed further below) can go a long way toward ensuring adequate data minimization and proportionality.

c. **Purpose Specification Principle—***This principle requires organizations to specify the purpose for data processing and avoid processing data incompatible with the stated purpose. This principle can be in tension with the general-purpose nature of many genAI models; developers may struggle to foresee and properly disclose with specificity in advance all the possible beneficial uses to which they or deployers may ultimately be able to apply the model.*

Purpose specification for genAI model training: Lawmakers and regulators should recognize the inherently broad purpose of training a general-purpose genAI model, as developers are training the model to respond to different commands or prompts and generate a range of potential outputs. General-purpose genAI models are intended to be deployed for a wide range of applications and many of these applications will be unknown at the time of development. Open-source models present even more uncertainty because, despite documentation detailing the kinds of tasks a model can perform,[33] details on usage are entirely with the deployer of the model, who likely has no contractual relationship with the developer.

The initial training of a genAI model cannot be seen as a singular, unrepeating stage of the development lifecycle; training is an iterative process and continues throughout the use of an AI system. Model 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 the robustness, accuracy, and security of the model.

The use of personal data for the purposes of training of general-purpose AI models should be recognized as a legitimate and permissible purpose, so long as other accountability measures and safeguards are reasonably and sufficiently implemented. For example, developers should be able to describe in plain language the stages of their model development process and the extent and purpose of personal data processing at each stage in relation to the model. Similarly, genAI deployers should provide clear explanations of how and why user-submitted personal data is used to operate their applications and whether user data will be used to train the model or improve the system in general.

To preserve the societal benefits[34] of general purpose genAI models, lawmakers and regulators should interpret purpose specification requirements flexibly during the development stage. They can require developers to provide sufficient transparency measures that indicate the range of applications or tasks that the model is well suited for, given the developer's resources and monitoring abilities (see additional discussion of transparency below). Such documentation should also, when possible and applicable, outline uses that the developer considers inappropriate (or prohibits as a condition of use), as well as use cases that the model may not be well suited for.

Developer and deployer distinctions: Using data to develop a genAI *model* should be treated as a separate purpose from using the data to develop, deploy and improve a specific *system* (i.e., an application that is built upon a genAI model). GenAI model developers should assess and set out the purpose of each stage of the development and training and establish, where necessary, what personal data is needed for that purpose. GenAI application deployers may proceed under separate and distinct processes and purposes that model developers may not have full insight into or control over. For this reason, application developers are the appropriate party to specify the purpose of processing in connection with the genAI application layer. That said, in some cases, a genAI model developer may build the model while simultaneously initiating work on potential applications. There should be an ability in such a case to use data across these processes.

d. Use Limitation Principle—*This principle seeks to prevent a "free-for-all" in an organization's use and re-use of personal data by requiring that the use of the personal data be limited to the stated purposes and reasonably aligned with the initially-stated purpose. While some observers may find this principle to be at odds with the development and use of genAI systems, steps can be taken to mitigate risks without foreclosing future technological breakthroughs associated with genAI models.*

While this principle remains important in the era of genAI, laws and regulations should allow organizations to articulate data use purposes that are sufficiently broad and flexible for the range of potentially useful applications for which the genAI model may be used. Purpose and use limitation principles should not be absolute.

There is a wide range of social and public benefits available from the use of data. Organizations should be allowed to re-use data to innovate in socially beneficial ways so long as they do so in an accountable and ethical manner by implementing appropriate safeguards to protect individuals' data privacy and fundamental rights. Data protection principles should be interpreted to limit unforeseen and harmful processing while enabling beneficial processing that is compatible with and does not undermine or negate the original purpose. Organizational accountability safeguards, including transparency and risk assessments that enable tailored mitigations, can ensure that new uses of data do not expose individuals to increased risks or adverse impacts.

In many contexts, it may be advisable that developers of genAI models build in controls to prevent users from creating detailed profiles of individuals, retrieving sensitive information about other users, or generating the likenesses of individuals without consent. Additionally, and perhaps more

crucially, providers and deployers of AI systems should implement appropriate controls at the output stage.

e. **Individual Rights—***Many data protection laws set out a number of individual rights such as the right to notice, access, correction, objection and in some cases erasure.*[35] *The provision of individual rights may require contextual interpretation given the novel nature of genAI models. For example, how can developers best provide notice to individuals when they rely on web scraped data and do not necessarily know at collection whether personal data was included in the collected set and do not have a direct relationship with individuals; and is it technically possible to erase personal data from a genAI model's memory (versus applying a filter to prevent the data from being included in further output)?*[36]

Notice: 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. 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, a deployer client of the model 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 in relation to personal data they process during their particular deployment of a genAI model.

Where appropriate and practicable, individuals should be able to request that their genAI input prompts and output responses not be included in model improvement and fine-tuning, especially if these include personal or sensitive data. Some organizations can impose controls or limits on their model's "chat memory" by default or offer individuals the ability to instruct the genAI system to remember or delete its memory of certain data they've inputted into the chat.

Some exceptions to providing notice should apply where data is not collected directly from individuals (e.g., in cases where data is collected via web scraping or web crawling). Datasets used for training genAI are vast, may sometimes be unstructured, and may in some cases include personal data only incidentally. A blanket requirement to identify individuals for notification purposes in web-scraped data would require large-scale additional processing purely for notification purposes. The effort of the organization in informing the individual must in each case be contextually balanced against the harm to the individual's rights should they not receive notice or an opportunity to exercise individual rights. Thus, organizations should be able to show that the required effort on a model developer's part to attempt to identify and subsequently provide the relevant privacy information to each individual whose data has been collected through web scraping meets the disproportionate effort exception to providing notice. In place of individual disclosures, organizations should be able to fulfil 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 model. It should be noted that the legitimate interests legal basis gives individuals the right to object to the processing of their data (for reasons relating to their particular situation).

Erasure: Lawmakers and regulators should consider that, in the case of web scraped data that is used during training but not catalogued or further filtered to identify personal data, it may be

unreasonable for a developer to respond to requests for erasure. Developers may be able to apply alternative measures, such as output filters, to satisfy an individual's erasure request.

Regulators should consider the full spectrum of compliance requirements organizations must meet to protect individual rights. 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. Laws and regulations 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.

Additionally, regulators should consider legal requirements outside of data protection, such as in banking, housing, education, health, civil rights, intellectual property, and contract laws, and how these requirements may impact the processing of personal data in the development and deployment of genAI. For example, employment, housing, and consumer protection laws across jurisdictions are increasingly mandating fairness requirements and prohibiting algorithmic discrimination. CIPL recommends close regulatory cooperation to ensure a unified and consistent approach across obligations that may exist in a single jurisdiction.

Objection: Importantly, under the GDPR, when organizations rely on the legitimate interest legal basis for processing, individuals have a right to object to processing. This provides individuals with an important level of control over their personal data. Therefore, organizations that rely on the legitimate interests legal basis should allow individuals to object to the use of their personal data for model operation, development, and improvement at any time in and accessible manner, and cease processing, unless the organization can demonstrate compelling interests that override the reasons for the objection.

f. Transparency—*This principle requires organizations to inform individuals about the collection and uses of their personal data, as well as certain processing activities, such as categories of personal data processed. This principle can also manifest in requirements to explain automated decision-making. Finally, transparency principles generally require organizations to inform individuals about their individual rights and how they can assert them vis-a-vis the organization's services. Transparency principles can be challenging in the genAI context because developers may limit disclosure of data sources due to trade secret purposes and because of the general-purpose nature of models, which may make it impossible for developers to enumerate all potential beneficial uses in advance.*

Transparency in the context of genAI models should be contextually appropriate, while also fulfilling transparency requirements under applicable laws and regulations. Organizations should make it possible for individuals to understand how their data is being used and transparency measures should enable individuals to reasonably exercise their privacy and data protection rights (e.g., right to be informed, right to object to processing, right to restrict processing, right to obtain rectification or erasure, etc.). However, the ability for organizations to satisfy individual rights requests may be dependent on the context, and the purpose and intended use of the model. Furthermore, the level of detail provided by transparency measures must be proportionate to the

risk posed by the processing, and organizations should recognize that the greater the risk posed by the processing, the higher the level of transparency that should be offered to individuals. Finally, transparency should not come at the expense of other important factors or principles, such as usability, functionality, and data security principles, or create additional burdens for users.

In all instances, the level of transparency should be balanced not only with the need to protect intellectual property rights, copyright, confidential information and trade secrets, but also the vulnerabilities of genAI systems and the potential net societal benefit that may outweigh individuals' rights. Risk assessments should, in most cases, be required to help organizations properly weigh these considerations. Organizations should also consider transparency in the wider sense, beyond individuals and users—to regulators, auditors, and red-team experts.

As described in CIPL's recent publication, *Building Accountable AI Programs: Mapping Emerging Best Practices to the CIPL Accountability Framework*, many genAI developers have published explanatory documents (e.g., model or system cards and technical reports) to provide transparency about genAI models. These documents can provide information such as how the model was built, how it was evaluated and what mitigations were applied, how it works, a summary of the types of data it was trained on, its intended use cases and contexts, key limitations, and performance metrics. Where possible and appropriate, such documents should also describe categories of personal data used in model training, including metadata on its key characteristics (e.g., what types of data are included in the dataset, where and how the data was collected, and which demographic groups are represented within it). They should also disclose information regarding what measures were taken to minimize reasonably foreseeable risks.

Laws and regulations should carefully balance transparency and data minimization principles. For example, any requirements to disclose details about the personal data contained in genAI training data could require developers to index the training data and take other measures that might be in tension with data minimization and therefore inadvertently increase privacy risks.

g. Organizational Accountability—*Data protection laws often integrate principles of accountability requiring organizations to implement demonstrable technical, contractual, and organizational measures to comply with various applicable data protection requirements. Legal and regulatory requirements incentivizing accountability give organizations a compelling motivation to invest in robust organizational accountability measures and programs, even as the environment for genAI development and deployment remains competitive and fast-moving.*

Responsible genAI development and deployment will always require organizations to carefully balance various rights, freedoms, and interests (including organizational, societal, and individual interests). Balancing tests, risk assessments, and mitigation measures are best demonstrated through organizational accountability measures that are meaningful, well-documented, and regularly updated. Just as laws and regulations require and incentivize organizations to develop and maintain robust privacy programs, they should also require comprehensive and risk-based AI programs. At the same time, organizations should proactively invest in such programs to

continually improve and evolve their controls and best practices and maintain a culture of responsible genAI development that functions throughout the entirety of the organization.

Responsible genAI development and deployment that meets data protection, accuracy, fairness, transparency, and security requirements, among other requirements, is a continuous journey. Lawmakers and regulators should encourage, facilitate, and reward organizational accountability in genAI development and deployment.

f. Cross-border Data Transfers—*Increasingly, nations are placing restrictions on certain cross-border data transfers. This may impede the development of accurate and fair genAI models as well as have other unintended negative consequences on beneficial genAI use.*

Lawmakers and regulators should consider the importance of cross-border data transfers for genAI model development: data flows ensure that model developers have access to sufficiently rich and diverse datasets to ensure the quality and fairness of system outputs and enable cross-border collaboration on beneficial research. Restrictions on such data flows may have unintended negative consequences for fairness and accuracy and impede beneficial research. Accountability measures, such as certification to cross-border privacy frameworks and privacy-enhancing technologies, can enable secure cross-border data transfers.

III. Conclusion

CIPL's practical and risk-based approach to privacy and data protection principles and how they apply to genAI seeks to promote innovation while ensuring robust protections for individual rights. While the emergence of genAI systems presents challenges to privacy rights and data protection, there is sufficient flexibility in data protection laws to protect individual rights while enabling beneficial and responsible uses of genAI systems. Open dialogue, cross-stakeholder engagement, including through regulatory sandboxes, and further research will remain central to the realization of responsible and accountable generative AI.

Endnotes

[1] Ina Fried, "OpenAI says ChatGPT usage has doubled since last year," *Axios*, 29 Aug. 2024, available at https://www.axios.com/2024/08/29/openai-chatgpt-200-million-weekly-active-users.

[2] Tiernan Ray, "Microsoft has over a million paying Github Copilot users: CEO Nadella," *ZDNet*, 25 Oct. 2023, available at https://www.zdnet.com/article/microsoft-has-over-a-million-paying-github-copilot-users-ceo-nadella/.

[3] McKinsey & Company, "The state of AI in early 2024: Gen AI adoption spikes and starts to generate value," 30 May 2024, available at https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai#/)(See Exhibits 1 & 3).

[4] IBM Research, "What are foundation models?" last updated 9 May 2022, available at https://research.ibm.com/blog/what-are-foundation-models.

[5] David Gewirtz, 'How does ChatGPT actually work?" *ZDNet*, 7 June 2024, available at https://www.zdnet.com/article/how-does-chatgpt-work/.

[6] Adam Zewe, "Explained: Generative AI," *MIT News*, 9 Nov. 2023, available at https://news.mit.edu/2023/explained-generative-ai-1109.

[7] Josh Hjelmstad, "Generative AI is Transforming Healthcare: Two Real-Life Success Stories," *AVIA Health*, 22 Aug. 2024, available at https://aviahealth.com/insights/generative-ai-transforming-healthcare-two-real-life-success-stories/.

[8] Zewe, supra footnote 6.

[9] The Hamburg Commissioner for Data protection and freedom of information, "Discussion Paper: Large Language Models and Personal Data," 15 July, 2024, available at https://datenschutz-hamburg.de/fileadmin/user_upload/HmbBfDI/Datenschutz/Informationen/240715_Discussion_Paper_Hamburg_DPA_KI_Models.pdf.

[10] In the US, a number of lawsuits have been filed against some organizations with genAI applications, alleging privacy violations; these are ongoing at the time of this paper: Andersen v. Stability AI; Doe v. GitHub, Inc.; Getty Images v. Stability AI; Leovy v. Google (see here for case citations: BakerHostetler, "Case Tracker: Artificial Intelligence, Copyrights and Class Actions," available at https://www.bakerlaw.com/services/artificial-intelligence-ai/case-tracker-artificial-intelligence-copyrights-and-class-actions/).

[11] This paper it is not meant to be a comprehensive discussion of all data protection principles and their application to genAI systems. Further, these concepts are sometimes named differently across various laws and frameworks (such as the OECD Privacy Guidelines and APEC Privacy Framework), and the legal ramifications will vary from one jurisdiction to another depending on the specificities of the law.

[12] Under the GDPR, the use limitation principle is included in Article 5, "Principles relating to processing of personal data". Where purpose specification requires organizations to specify the purpose for which they are processing data, use limitation requires them to use data only for the stated purpose or a compatible one.

[13] *See* CIPL Report, "Building Accountable AI Programs: Mapping Emerging Best Practices to the CIPL Accountability Framework," Feb. 2024. *See also* CIPL, "Ten Recommendations for Global AI Regulation", Oct. 2023. *See also* CIPL Report, "Hard Issues and Practical Solutions," Feb. 2020. *See also* CIPL Report, "Artificial Intelligence and Data Protection in Tension," Oct. 2018. All reports available at https://www.informationpolicycentre.com/cipl-white-papers.html.

[14] *See* CIPL "Response to CNIL How-To Sheets on the Development of Artificial Intelligence Systems", 1 Oct. 2024. *See also* CIPL "Compilation of Responses to UK ICO Generative AI Consultations," 20 Sept. 2024. *See also* CIPL "Response to the National Institute of Standards and Technology (NIST)'s Request for Comment on the Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile," 31 May 2024. All responses available at, https://www.informationpolicycentre.com/public-consultations.html.

[15] As it is called in the UK GDPR and EU GDPR.

[16] *See*, for example, Office of the Australian Information Commissioner, "Guidance on privacy and developing and training generative AI models," last updated 23 Oct. 2024, available at https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/guidance-on-privacy-and-developing-and-training-

generative-ai-models ("Where sensitive information is inadvertently collected without consent, it will generally need to be destroyed or deleted from a dataset.").

See also Michihiro Nishi, Clifford Chance, "Japanese Law Issues Surrounding Generative AI," Oct. 2023, available at https://www.cliffordchance.com/insights/resources/blogs/talking-tech/en/articles/2023/10/Japanese-Law-Issues-Surrounding-Generative-AI.html (Discussing the Personal Information Protection Commission's requirement that OpenAI "take necessary measures such that sensitive information is not collected in the first place and/or if such sensitive data is obtained, to take measures to remove the sensitive information as far as possible from the dataset immediately after collection, and remove or anonymise the sensitive personal data before converting the collected data into a dataset for learning.").

[17] Regulation 2024/1689, Article 10(5).

[18] CIPL, "Privacy-Enhancing and Privacy-Preserving Technologies: Understanding the Role of PETs and PPTs in the Digital Age," 12 Dec. 2023, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-understanding-pets-and-ppts-dec2023.pdf.

[19] For example, GDPR, Article 5 "Personal data shall be: (a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency');" New Zealand Privacy Act, Section 22, Information privacy principle 4(b)(i) "An agency may collect personal information only by a means that, in the circumstances of the case (particularly in circumstances where personal information is being collected from children or young persons), is fair . . ;" Canada Personal Information Protection and Electronic Documents Act, Section 4.4 " . . . Information shall be collected by fair and lawful means."

[20] *See* Dr. Richard Fletcher, "How many news websites block AI crawlers?" 22 Feb. 2024, available at https://reutersinstitute.politics.ox.ac.uk/how-many-news-websites-block-ai-crawlers.

[21] EDPB, Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, available at https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2024/guidelines-12024-processing-personal-data-based_en. *See also* ICO, "How do we apply legitimate interests in practice," available at https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/legitimate-interests/how-do-we-apply-legitimate-interests-in-practice/ (last visited 26 Nov. 2024).

[22] Case C-621/22 *Koninklijke Nederlandse Lawn Tennisbond v Autoriteit Persoonsgegevens* [2024] ECLI:EU:C:2024:857, paras 48-49 & 57.

[23] Joined Cases C-26/22 and C-64/22 *SCHUFA holding (Discharge from remaining debts)* [2023] ECLI:EU:C:2023:958, para 83.

[24] Regulation 2024/1689, Article 10(5). However, Article 10(5) only applies to AI models considered "high-risk" under the AI Act and there remains a legal gap concerning the processing of special category data to prevent bias in non-high-risk systems.

[25] Case C-136/17 *GC and Others* [2019] ECLI:EU:C:2019:773, para 47.

[26] GDPR Article 9 prohibits the processing of special category data (sensitive data) unless one of 10 exceptions applies, including individual consent or the processing is related to data that an individual "manifestly made public."

[27] Case C-252/21 *Meta Platforms Inc and Others v. Bundeskartellamt* [2023] ECLI:EU:C:2023:537, para 77.

[28] *Id.* at paras 77-85.

[29] GDPR Article 9(2)(g).

[30] *Id.* at Article 9(2)(j).

[31] Google Search Central Documentation, "Introduction to robots.txt," available at https://developers.google.com/search/docs/crawling-indexing/robots/intro (last visited 26 Nov. 2024).

[32] Some research shows that genAI language models can produce diminished outputs over time when they rely on synthetic data for training purposes. *See* Yanzhu Guo et al., "The Curious Decline of Linguistic Diversity: Training Language Models on Synthetic Text," last revised 16 Apr. 2024, available at https://arxiv.org/pdf/2311.09807.

[33] E.g., text summarization, question & answer, or image classification. *See* Hugging Face Open-Source Model Library, available at https://huggingface.co/models.

[34] GenAI systems have turbo-charged access to a wide range of tools and tasks, including, real-time translation tools, medical diagnosis, cybersecurity, compliance tools, and personalized tutors.

[35] For example, Chapter 3, GDPR, Rights of the data subject (Article 15, Right of access; Article 16, Right of rectification; Article 17, Right to erasure; Article 18, Right to restriction of processing). See also Chapter III, LGPD, Data Subject's Rights (Article 18, Right to obtain (i) confirmation of the existence of the processing; (ii) access to the data; (iii) correction of incomplete, inaccurate or outdated data; (iv) anonymization, blocking or erasure of unnecessary or excessive data or data processed in noncompliance with the provisions of this Law; (v) portability of data to another service . . .).

[36] Tensions regarding the provision of individual data protection rights are common in emerging technologies and it's important to consider how previous tensions have been addressed. For example, the immutable nature of blockchain and the distributed ledger means in principle that all transactions are recorded forever, and deletion is not an option. In analyzing how this may impact the right to erasure, the French data protection authority (the CNIL) acknowledged that some encryption techniques, coupled with key destruction, can potentially be considered erasure even if erasure in a literal sense is not possible. *See* CIPL Discussion Paper, "Digital Assets and Privacy," Jan. 2023, pg. 21, available at https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_discussion_paper_on_digital_assets_and_privacy__19_jan_2023_.pdf.