GPA AIWG

Interim report on the work conducted by the AIWG members on generative AI systems

SUMMARY

INTRODUCTION

Following the adoption of the Resolution on Generative Artificial Intelligence Systems during the 45th Closed Session of the Global Privacy Assembly on October 2023, it was decided to present an interim report on the work conducted by the GPA AIWG members on generative AI systems during the 46th Global Privacy Assembly.

In order to complete this report and as part of its work program for 2024, a survey was circulated to the members of the GPA AIWG on the work they are conducting on generative AI systems, and notably how they implement the GPA resolution adopted in 2023.

Answers were received from the following members:

- Côte d'Ivoire (ARTCI)
- Mexico:
  - INFOCDMX
  - INFOEM
- Canada:
  - OIPC (Columbia)
  - OPC (Federal)
  - OIPC (Ontario)
- Switzerland (FDPIC)
- Kenya (ODPC)
- New Zealand (OPC)
- Japan (PPC)
- Uruguay (URCDP)
- Korea (PIPC)
- Guernsey (ODPA)
- Argentina (CPDP)
- Hong Kong (China)
- Israel (PPA)

I - AI REGULATIONS

As recalled in the GPA resolutions on generative AI systems, data protection and privacy principles and current laws, including data protection and privacy laws, bills, statutes and regulations, apply to generative IA products and services, even as different jurisdictions continue to develop AI-specific laws and policies. Data protection authority have jurisdiction over generative AI systems insofar as the processing of personal data is involved.

Only a few jurisdictions have adopted or are currently developing specific laws, regulations or rules on generative AI systems.

In the EU, the Artificial Intelligence Act (AI Act) has been published in the Official Journal of the European Union on July 12th, 2024 and will gradually come into force as of 1 August 2024. It is the world's first general legislation on artificial intelligence. It aims to provide a framework for the development, placing on the market and use of artificial intelligence (AI) systems, which may pose risks to health, safety or fundamental rights. The AI Act follows a risk-based approach by classifying AI systems into four levels:

- Unacceptable risk: the AI Act prohibits a limited set of practices that are deemed contrary to the values and fundamental rights of the EU.
- High risk: The AI Act defines AI systems as high-risk where they may affect the safety of individuals or their fundamental rights, which justifies their development being subject to enhanced requirements (conformity assessments, technical documentation, risk management mechanisms).
- Specific transparency risk: the AI Act imposes specific transparency obligations on AI systems, in particular where there is a clear risk of manipulation.
- Minimal risk: for all other AI systems, the AI Act does not include a specific obligation. These are the vast majority of AI systems currently in use in the EU or likely to be used in the EU according to the European Commission.

The EU AI Act provides a framework for a new category of general-purpose AI models, in particular in the field of generative AI. These models are defined by their ability to serve a large number of tasks, making them difficult to classify them in the previous categories. For this category, the AI Act provides for several levels of obligations, ranging from minimum transparency and documentation measures to an in-depth assessment and the implementation of systemic risk mitigation measures that some of these models might entail, in particular because of their power: risks of major incidents, misuse to launch cyberattacks, the spread of harmful biases (e.g. ethnicity or gender) and discriminatory effects against certain persons, etc.

In Canada, even if there are no current laws specific to artificial intelligence, several Bills are considered at the federal and States level. At the federal level, the Canadian House

of Commons is currently considering Bill C-27 (the *Digital Charter Implementation Act, 2022*), which would (among other things) enact the *Artificial Intelligence and Data Act* (AIDA). In the State of Ontario, Bill 194, Strengthening Cyber Security and Building Trust in the Public Sector Act, 2024 - Legislative Assembly of Ontario (ola.org) was introduced in May 2024, and addresses cyber security and artificial intelligence systems at public sector entities. Moreover, and always in Ontario, Bill 149, Working for Workers Four Act, 2024 - Legislative Assembly of Ontario (ola.org) Bill includes amendments to the rules regarding job postings in the Employment Standards Act. The amendments include a requirement that employers disclose the use of AI in the hiring process. Also, Innovation, Science and Economic Development Canada (Canada's Ministry of Industry) has also released a voluntary code of conduct on the responsible development and management of advanced generative AI systems. This is not, however, an enforceable rule.

In Korea, Ministry of Science & ICT has been making effort for enactment of an AI law that can strike a balance between growth and trust. Many bills are currently proposed before the National Assembly for the development of trustworthy AI in Korea. Highlights of the major bills regarding AI are as follows:

- Purpose: Stipulating the basic details to support the development of reliable AI
- High-risk AI Systems: Clarifying the definitions of high-risk AI and setting out accountability of AI service providers dealing with high-risk AI systems
- Launching a National AI Commission (TBD): Establishing a presidential commission to realize AI society and deliberate on the matters associated with AI, such as building trust in the AI industry
- Master plan: Stipulating the need for establishing AI policy directions and relevant strategies on a regular basis
- Ethics of AI: Setting out AI ethics to ensure accountability and trustworthiness associated with AI

II - PUBLICATION OF GUIDANCE ON THE APPLICATION OF DATA PROTECTION AND PRIVACY PRINCIPLES TO GENERATIVE AI SYSTEMS BY MEMBERS OF THE AIWG

1. Canada (OPC and provincial and territorial commissions)

In December 2023, Canada's federal, provincial and territorial privacy commissions jointly published "Principles for responsible, trustworthy and privacy-protective generative AI technologies." These principles are targeted at organizations developing, providing or using generative AI systems to help them apply key Canadian privacy principles. It includes sections on:

- Legal authority and consent (i.e. lawful basis)
- Appropriate purposes
- Necessity and proportionality
- Openness (i.e. transparency)
- Accountability
- Individual access (i.e. rights of data subjects)
- Limiting collection, use and disclosure
- Accuracy
- Safeguards

The OPC has also led the drafting of an International Working Group on Data Protection in Technology (IWGDPT, "Berlin Group") paper on Large Language Models, which will speak to privacy principles broadly rather than any specific principle, that should be published in second half 2024.

Finally, the OPC has identified "Addressing the privacy impacts of the fast-moving pace of technological advancements, especially in the world of artificial intelligence (AI) and generative AI" as a strategic priority for 2024-2027. To this end, the OPC intends to publish additional documents on AI over the next three years (likely focused on practical implementation of established principles).

2. EDPS

The EDPS released first orientations for ensuring data protection compliance when using generative AI systems in June 2024. The aim of this publication is to provide practical advice and instructions to EU institutions, bodies, offices and agencies on the processing of personal data when using generative AI systems. This paper includes notably sections on lawfulness, data minimisation, data accuracy, transparency, the exercise of individual rights, automated decision making, fair processing and avoiding bias, and data security.

More detailed guidance is planned for the future.

3. France (CNIL)

In July 2024, the CNIL published a Q&A on the use of generative AI systems offering first answers to organisations that consider deploying or using generative AI systems on the following items: the benefits and risks of generative AI systems, what approaches are available today to use generative AI and how to choose a generative AI system, deployment method, implementation and management, training of end-users, governance, compliance with GDPR and the AI Act.

As regards compliance with the GDPR, the CNIL has published AI how-to-sheets for the creation of data bases used to train AI systems, which involve personal data. Following topics are covered by the how-to-sheets: applicable legal regime, purpose, legal qualification of AI system providers, lawfulness of the data processing (legal basis, instances of re-use of data, DPIA) privacy by-design. Following topics are currently opened for consultation: legal basis of legitimate interest (also with a focus on open source models and on web scraping), informing data subjects, exercise of data subjects' rights, annotating data, and security.

4. Guernsey (ODPA)

In May 2023, the ODPA published an overarching piece of guidance in respect of AI which touches all the key principles listed in the GPA resolutions, and underscores that local data protection law applies to AI systems that use personal data.

5. Hong Kong (PCPD)

The PCPD published the "Guidance on the Development and Use of AI" in 2021. While this Guidance does not pinpoint any types of AI, the ethical principles and recommendations contained therein are applicable to organisations which develop or use generative AI systems that involve the use of personal data.

The PCPD also published the "AI: Model Personal Data Protection Framework ("2024 Framework") in June 2024. The 2024 Framework builds upon the 2021 Guidance and targets organisations which procure, implement and use any type of AI systems. While the 2024 Framework does not focus specifically on generative AI either, it contains more recommendations applicable to the processing of personal data in the customisation and use of generative AI by organisations.

While both the 2021 Guidance and the 2024 Framework provide recommendations on compliance with the Personal Data (Privacy) Ordinance (PDPO) in the context of AI, the PDPO is a piece of legislation that is technology-neutral and principle-based. The provisions and principles on personal data protection therein apply equitably to any technical means of collecting, using, storing, retaining and transferring personal data, including generative AI. PDPO's regulation covers six personal data protection principles (i.e. purpose and means of personal data collection; accuracy, storage and retention of data; use of data; data security; transparency of data policies; and data access and correction), so as to ensure that the entire process of the handling of personal data is subject to legal safeguards.

6. Israel (PPA)

The PPA wrote the chapters regarding privacy and data protection risks and means to mitigate them in the comprehensive "Policy on Artificial Intelligence Regulation and Ethics", published in December 2023 by the Ministry of Innovation, Science and Technology along with the Ministry of Justice.

The PPA has also issued:

- An opinion on Privacy and Data Security in Deepfake Technologies
- An opinion about social scoring and data protection (only in Hebrew)
- Guidelines about transparency including transparency in AI (only in Hebrew)

In addition, the PPA intends to publish guidelines regarding the application of the Privacy Protection Law to AI systems.

7. Japan (PPC)

On June 2023, the PPC has issued notices on the use of generative AI services (available only in Japanese). This is a concise guidance for businesses handling personal information, administrative entities, and general users to comply with the Act on Protection of Personal Information (APPI) when using generative AI services rather than those developing, providing, and/or deploying generative AI product or services.

On June 2023, the PPC also issued a regulatory notice to OpenAI. This notice is issued as an administrative instruction in accordance with the provision of the APPI. The outline of the regulatory notice to OpenAI, developing and providing a generative AI service called "ChatGPT" is available (only in Japanese). It requested Open AI to do as follows:

- To take safeguard measures to prevent or mitigate risks and to ensure proper handling of sensitive personal information, in particular when the company

collects or processes such information for training the model of the generative AI (ChatGPT).
- To disclose a purpose of use of personal information for ChatGPT in Japanese (as OpenAI had disclosed the purpose of use only in English at that time).

The notice was originally addressed to OpenAI, but it serves as a reference for other developers, providers and deployers of generative AI services as well.

The documents refer to the following principles:

- Lawful basis for processing,
- Purpose specification and use limitation,
- Data minimisation,
- Accuracy,
- Transparency,
- Privacy by design and by default,
- Rights of data subjects.

8. Kenya (ODPC)

The ODPC advises data handlers with Generative AI systems to incorporate and uphold data protection principles in line with existing best practice and international standards, notably the GPA resolution on Generative AI.

9. New Zealand (OPC)

In June 2023, the OPC published a statement on privacy and generative AI in which it expects that agencies considering implementing a generative AI tool will have senior leadership approval, review whether a generative AI tool is necessary and proportionate, conduct a privacy impact assessment, be transparent, engage with Māori, develop procedures about accuracy and access by individuals, ensure human review prior to acting, ensure that personal information is not retained or disclosed by the generative AI tool.

This statement was followed by the publication of a guidance on Artificial Intelligence and the Information Privacy Principles in September 2023, which covers all principles listed in the GPA resolution except rights of data subjects.

10. Korea (PIPC)

In August 2023, the PIPC has released policy direction for "Safe Use of Personal Information in the Age of AI". The main points of the policy direction are as follows:

- The PIPC will take a “principle based” approach to regulation, rather than a prescriptive “rule-based” approach.
- The AI Policy describes in detail the principles and standards applicable to the processing of personal data for each stage of AI lifecycle (Design and planning - Data collection - Model building and training - Provision of AI Services).
- The PIPC plans or has already publish further guidelines for more specified needs
    - Guideline on pseudonymization of unstructured data (released Feb ‘24)
    - Processing of Biometric Data: exploration of policy directions
    - Guideline on Processing Publicly Available Data (released July ‘24)
    - Usage of data captured by mobile data processing devices for images: exploration of policy directions
    - Guideline on Processing Personal Data to Ensure Transparency for AI Development and Services
    - Guideline regarding Usage of Synthetic Data

The documents address all privacy principles listed in the GPA resolution on generative AI systems.

11. United Kingdom (ICO)

The ICO has launched a consultation series on the application of data protection law to the development and use of generative AI models, notably on lawfulness, purpose limitation, data accuracy, exercising data subjects’ rights

III - SUPPORTING MEASURES PUT IN PLACE BY GPA AIWG MEMBERS TOWARDS PUBLIC OR PRIVATE ENTITIES SUCH AS REGULATORY SANDBOX

Most of the supporting measures put in place by respondents to the survey are broadly designed to accompany the implementation of innovative projects and can be used for the development or use of generative AI systems and.

Some respondents have implemented specific measures on generative AI, or on AI systems more broadly:

- The OIPC of Ontario has initiated non-legally binding consultations with few public institutions, guiding them how to deploy, and what to take into consideration, when deploying generative AI solutions in their services. These consultations are mostly on general guidelines, best practices, etc.
- The CNIL (France) supports 8 AI innovative projects aimed at improving public services by providing them personalised accompaniment over several months, and has also selected 3 companies with strong potential scale-up AI companies to benefit from enhanced support over several months.

Main feedbacks from these supporting measures:

- It can feed the doctrine of the Regulator by providing concrete use cases and examples to reflect and work on,
- The continued dialogue between regulators and organizations on a long period contributes to improve regulators' knowledge of the practices of sectors concerned and to better understand constraints that entities face in implementing the law,
- This kind of cooperation allows also to have a clearer idea of the maturity level of a sector in terms of compliances,
- Entities often struggle to understand how to apply existing law to emergent technologies,
- Entities often struggle to understand the technologies themselves,
- Lack of clear legislation on many uses of AI technologies can make the process more challenging for regulated entities.

IV – ENFORCEMENT ACTIONS UNDERTAKEN BY AIWG MEMBERS REGARDING GENERATIVE AI SYSTEMS

Several respondents have initiated enforcement actions with respect to the practices of generative AI systems, often through a national or regional framework coordinated:

- The Ibero-American Network of Data Protection Authorities (RIPD) has initiated a coordinated enforcement action regarding ChatGPT in August 2023.

- Several EU data protection supervisory authorities are investigating ChatGPT and coordinated within the EDPB: through the one-stop-shop mechanism for cross-border processing carried out by OpenAI from February 15th, 2024 (date from which it had a single establishment in the EU), and trough the Task Force on ChatGPT for processing operations carried out until February 15th, 2024 and that concern possible infringements of non-continuing or non-continuous nature. A report on the work undertaken by this task force was released in May 2024.

- In Canada, the OPC and provincial authorities from British Columbia, Alberta and Quebec are jointly investigating ChatGPT, following a complaint alleging the collection, use and disclosure of personal information without consent.

- In Korea, the PIPC imposed a fine of KRW 3.6 million (around 2398 euros) on Open AI, for failing to report data breach related to payment system and also announced a list of domestic cases of non-compliance with the Personal Information Protection Act relevant to transparency, legal basis for processing (lack of consent), lack of transparency in consignment, and absence of parental consent for children under the age of 14. In March 2024, the PIPC established enhanced protection measures based on the pre-emptive inspection of Six businesses that develop or deploy Large Language Models (LLMs) or provide AI services powered by LLMs, including OpenAI, Google, Microsoft, Meta, Naver and Wrtn Technologies and recommended to implement improvements:
    - Enhanced protection for personal data such as unique identification data, credit card information, etc. in training data
    - Notification of the fact and purpose of human review of user input data and ensuring the right to choose regarding data usage
    - Setup of a response system to promptly address service vulnerabilities when discovered.

Some respondents flagged that most of the complaints lodged regarding AI systems were about facial recognition systems using AI, and not necessarily about generative AI.