(2024) 36 SAcLJ 307
(Published on e-First 19 July 2024)

GENERATIVE ARTIFICIAL INTELLIGENCE

The Protection of Personal Data and Countering False Narratives About the Person

Generative artificial intelligence ("Gen AI") has rapidly become ubiquitous on online platform services, from chatbots and virtual assistants to search engines and social media. This generated concerns over potentially harmful effects from its use in both social and professional settings, including the added threats to personal data privacy and accuracy of personal information. In this article, the author will explain how Gen AI operates and why it gives rise to these issues, examine the policy and law relating to Gen AI, both existent and anticipated, and suggest possible solutions to the problems in the form of legal and non-legal measures.

Warren B **CHIK**
LLB (National University of Singapore), LLM (Tulane), LLM (UCL); Associate Professor of Law, Yong Pung How School of Law, Singapore Management University; Deputy Director, Centre for AI and Data Governance.

I. Introduction

1 The general apprehension over loss of data privacy and the potential misuse of personal data through the use of "generative artificial intelligence" ("Gen AI") is compelling because of the immense success of ChatGPT, developed by OpenAI, within a short period of time. Furthermore, ChatGPT precipitated the launch of a more diverse range of Gen AI models by other organisations (and OpenAI itself) in an apparent race to capture a significant market share of users and in the case of paying models, subscribers. The speed of these developments is itself worrying as it bypasses the opportunity to study its repercussions and to prepare a proper response prior to operationalisation.

2 Gen AI are mainly owned and operated by private entities ("Gen AI organisations") that have an incentive to protect the informational privacy relating to the training and development of their Gen AI models. Also, there are (as at the time of writing) no laws in any country or region that specifically govern and control the development and deployment of AI in general, although there are a few AI laws which are at various stages of the legislative process. Hence, there are genuine concerns over how

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

these systems will operate in the absence of such laws and regulations, such as the lack of transparency over the type of algorithm and training data used to build the Gen AI. There also appears to be a lack of accountability over the extent that Gen AI makes "independent" decisions and produces content without human oversight, and the possible effects on society. Further questions also arise relating to the ethics of its use and the appropriate responsibility for Gen AI-created content.

3 This article will mainly focus on personal data protection concerns and, to the extent that it is relevant, false information about an individual. In Part II of this article, the author will provide an overview of the rise of Gen AI. Prominent examples of Gen AI, their purpose and uses, and the challenges that arise from their development and operation will be explained. In particular, the problems that Gen AI pose to the informational veracity and the overall responsible management of personal data will be highlighted. This is important as it will be particularly relevant to the question of allocation of responsibility, the need (or otherwise) for regulation and the type of oversight over Gen AI organisations that should be put in place. Hence, the current and anticipated regulatory treatment of, and the global policy trend towards, Gen AI will be canvassed in Part III.

4 In Part IV, the author will consider whether the current laws relating to the protection of personal data and against its misuse are applicable to Gen AI systems. This includes how the current laws can be purposively interpreted to respond to Gen AI. In particular, the scope and understanding of what constitutes "personal data" for the purpose of protection will be revisited as a result of the impact that Gen AI can have on personal data privacy, directly or indirectly. The creation and dissemination of false information by Gen AI and the adequacy and appropriateness of the existing measures available to deal with it, particularly in relation to false narratives about the person, will also be examined. In the course of doing so, the author will provide some suggestions and recommendations, both operational and legal, to improve on the current system of governance over Gen AI in the Singapore context.

II. Short history on Gen AI, its functions and impact on society

A. Quick rise of Gen AI and race to launch

5 While most policy and lawmakers were distracted by the possible ethical concerns relating to the metaverse and the use of its related devices for several years preceding the steep rise in the use of Gen AI, it only took a few months after the launch of ChatGPT on 30 November 2022, with

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

its swift take-up rate and the effects it had (and could potentially have) on society and the economy, for the focus to quickly pivot to Gen AI as a category of AI of greater and more immediate concern.[1]

(1) What is Gen AI and how does it work?

6 Gen AI refers to machine learning models that can be trained with data (*ie*, "input") to create content. These models are trained on large datasets and can learn patterns and structures from them in order to speedily produce content (*ie*, "output") based on the training data when prompted. It differs from the more rudimentary predictive AI because of its ability to generate new and original content. The output is "realistic" in that it appears like it is created by a natural person; *eg*, large language models ("LLMs") produce a realistic human-like conversational language that is presented in response to instructions or questions from users. An example of a prominent category of Gen AI are generative adversarial networks.[2] Another category is that of variational autoencoders.[3]

7 Another important feature of Gen AI is that it works on probability in generating output; *ie*, the output is based on learned statistical distributions. It also does not "save and reuse" an output template, but rather, it generates a fresh output based on each user prompt and the output for each case may differ to some extent, even when given the same or similar prompt. The output may exhibit characteristics similar to the training data and possibly replicate it to some extent.

8 Aside from Gen AI which use LLM that produce text, which the general public is probably most familiar with, other types of machine learning models (or "foundation models" known as "FM") are capable of creating other forms of new content such as images (still and moving, realistic or abstract), music and speech and computer codes. As with LLM, the objective is to develop content based on user prompting, while

1 There are obstacles to the widespread access to and use of the metaverse due to the high cost of the technology required (*eg*, wearable devices). Also, the metaverse has yet to provide a seamless and immersive experience. In contrast, Gen AI like ChatGPT has free-to-use and paid options and provides immediate results.

2 See Antonia Creswell *et al*, "Generative Adversarial Networks: An Overview" (January 2018) 35(1) *IEEE Signal Processing Magazine* 53 and Kunfeng Wang *et al*, "Generative Adversarial Networks: Introduction and Outlook" (2017) 4(4) *IEEE/CAA Journal of Automatica Sinica* 588.

3 See Joseph Rocca, "Understanding Variational Autoencoders (VAEs)" *Towards Data Science* (24 September 2019) <https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73> (accessed 31 May 2024) and George Lawton, "What Is a Variational Autoencoder (VAE)" *TechTarget* (June 2023) <https://www.techtarget.com/searchenterpriseai/definition/variational-autoencoder-VAE> (accessed 31 May 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

mimicking the natural person in the manner of expression through the delivery and presentation of the output material. In this way, Gen AI can perform tasks in an organised fashion, prompted by a user, that is almost already fit for purpose and that a real person is likely to be receptive towards, given the *coherence of* the content that is *solicited by* the recipient.

9 The types of Gen AI and their purposes are varied and can give rise to greater nuances in the socioeconomic and political issues they raise. The focus of this article will be primarily on Gen AI producing text, images and sound – because they are currently the most prevalent for public use and mass consumption – and on data privacy issues (especially in relation to personal data) as well as data integrity (particularly veracity) concerns. It should be noted at the outset that the development of new Gen AI models is accelerating at breakneck pace, with one of the latest (as of the date of this publication) being the introduction of Sora,[4] a new text-to-video Gen AI model, on 16 February 2024 by OpenAI.

(2) Objective and purpose of Gen AI?

10 Like all AI, Gen AI primarily helps to perform tasks faster and more efficiently. Specifically in relation to LLMs, it can *appear* to the general user as the progeny of a search engine combined with an online encyclopaedia, like Google and Wikipedia respectively, in the way it responds to queries and presents results in a "customised" manner, resembling the output from a chatbot. Gen AI also fall within the category of "pull" technology, as it is the user that frames the request and determines when and how information is received.

11 However, it is too simplistic to treat Gen AI as a more advanced prototype or iteration of a cataloguing database and online information depository, because the type of technology and the nature of the data and presentation are quite different. Gen AI can also perform many other tasks like summarising texts and replicating a writing style.

12 Moreover, search engines and online encyclopaedias use different technical methods to create results and content through their own patented algorithms and content development processes respectively. The former does not create new content while the latter relies on third-party contributions, which makes them significantly divergent from Gen AI with different legal implications (although there are some common concerns). For example, there is the issue of allocation of responsibility and liability over Gen AI content, since the creation (and creator) of

4 See <https://openai.com/index/sora/> (accessed 4 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

content as well as the possession and use of source material differ between Gen AI on the one hand, and search engines and online encyclopaedias on the other.

13 Online information depositories present information, updated periodically (but not as frequently as Gen AI), in "permanent" form on their websites (*ie*, these do not deliver tailored or customised output to the same or similar keyword search). Information locator services typically present data in the form of a search engine results page, normally displaying hyperlinks to the source webpage and featuring snippets of its content, in a "ranking" format (and as sponsored or organic results). There is also a time lag in the presentation of "www" content, depending on the frequency at which the cache of a webpage is updated. Gen AI is much more personalised and original in content delivery and presentation. This is another important difference, which makes it more challenging to monitor and ensure compliance with personal data interests (and the enforcement of such rights), and intensifies veracity concerns surrounding the use of Gen AI in comparison. It will consequently have an impact on the effectiveness of existing laws on personal data protection and laws against the creation and dissemination of false information.

(3) How is Gen AI being used and notable Gen AI models launched?

14 The intense interest and use of Gen AI can be attributed to the release of ChatGPT[5] by OpenAI, an AI research and deployment company, for use by the general public.[6] It is an LLM GPT[7] chatbot that currently supports many languages including English, French, Spanish, German and Chinese. As noted at para 5 above, OpenAI released a demo version of ChatGPT on 30 November 2022, which quickly went viral and gained popularity for its usefulness and efficiency.[8]

15 Gen AI can be used for free or with paid premium services, and can be specifically customised and licenced for commercial use. It can be used in a "closed" or private environment, such as within an organisation or company, or in a network or conglomerate. It can also operate in an "open" or public environment, *ie*, offered for public use by anyone

5 Version 3.5.

6 However, nascent work on the idea and theory of Gen AI can be traced back to as early as the 1940s, according to the earliest available literature on the subject. See C E Shannon, "A Mathematical Theory of Communication" (July, October 1948) 27 *The Bell System Technical Journal* 379 at 623 and A M Turing, "I. – Computing Machinery and Intelligence" (October 1950) LIX(236) *Mind* 433.

7 "GPT" stands for "Generative Pre-trained Transformer".

8 OpenAI also created other Gen AI models like DALL-E for images, Whisper for speech transcription and Sora for video generation from text instruction.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

with Wi-Fi access. In some countries, including Singapore, even the government is considering the use of Gen AI for the delivery of public sector government services, including the provision of administrative and legal services.

16 A notable Gen AI launched by OpenAI other than ChatGPT is DALL-E,[9] which was released in early 2021 to create images from textual descriptions. It uses a combination of unsupervised learning and reinforcement learning from human feedback. ChatGPT and DALL-E are also examples of the increasing use of user feedback as training data to further improve output, which also have personal data concerns. We also see other advances in the text-to-image Gen AI models like Stable Diffusion,[10] released in 2022, which was incubated and developed by academic researchers. In tandem with text and image models, advancements were also made in the field of AI voice generation and text-to-speech capabilities.

17 It is noteworthy how Gen AI is integrated to improve existing services as well as to provide an alternative to the "traditional" forms of content delivery. Unsurprisingly, many of the big technology companies operating search engines and social media platforms are also improving their existing services by incorporating Gen AI into their processes. For example, BART[11] (owned by Meta Platforms Inc, formerly Facebook Inc) was developed to improve natural language processing tasks like summarising texts and language translation. Another example is BERT[12] (owned by Google LLC), the acronym for "Bidirectional Encoder Representations from Transformers", which was designed for natural language understanding tasks to provide language translation, understand the context of words in sentences and to answer questions.

18 Bing[13] (owned by Microsoft Corporation) was relaunched as an AI-powered search engine to produce improved and more accurate search results, and also to present search queries as complete answers to questions by summarising answers from different sources. In competition with ChatGPT, it provides a "chat experience" and customised content,

9 See <https://openai.com/index/dall-e-3/> (accessed 4 June 2024).
10 See <https://stability.ai/> (accessed 4 June 2024).
11 See <https://research.facebook.com/publications/bart-denoising-sequence-to-sequence-pre-training-for-natural-language-generation-translation-and-comprehension/> (accessed 4 June 2024).
12 See <https://research.google/pubs/bert-pre-training-of-deep-bidirectional-transformers-for-language-understanding/> (accessed 4 June 2024).
13 See <https://www.bing.com/> (accessed 4 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

such as fully-formed e-mail messages and travel itineraries, in answer to queries relating to those formats.[14]

B. Precision bias and "illusion of accuracy"

19 ChatGPT's rapid popularity can be attributed to the *free usage* of GPT-3 (now GPT-3.5, while offering an upgrade plan for access to GPT-4, DALL-E and other perks), its *wide release for public use*, the *breadth of its content* (which is still growing exponentially) and *how the output is presented*. The user experience is smooth and results are instantaneous. It can also perform many tasks, personal or professional, in a much shorter time. For example, it can draft school or work assignments that may otherwise take more time or cost the user a fee (if third-party professional services are used), recommend travel itineraries that obviate the need to refer to travel books (or the services of travel agents) and prepare e-mail responses (and even presentations).

20 With its easy-to-read (and copy) format and specific (*ie*, customised) responses to search queries, Gen AI models like ChatGPT will rapidly gain popularity and have the potential to become as regularly and commonly used as traditional search engines, online encyclopaedias and social media services. These may perhaps even have the disruptive effect of surpassing and partially replacing such services. Such Gen AI models have already provided the impetus for many of these existing platforms to upgrade their services with Gen AI capabilities to remain competitive. As noted above, it has accelerated the development and operationalisation of such technology by some of the bigger companies that have already been researching and developing the technology, as seen with Google's BERT and Meta's BART.

21 Because information is presented in human-readable form and also simulates the natural person in expression, even in different languages, the output can be highly convincing as reliable information. This is especially so for the user who solicited the information and who is consequently more inclined to reuse it. It is also difficult for the average user, especially one not trained to use Gen AI responsibly, to discern if the information is accurate, to be aware of the source of the data provided and to know how results are generated. For information that is shared "downstream" by users to others, it becomes even harder for those persons

14 Yusuf Mehdi, "Reinventing Search with a New AI-powered Microsoft Bing and Edge, Your Copilot for the Web" *Microsoft* (7 February 2023) <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/> (accessed 31 May 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

to discern the source of the data (*ie*, whether it was human-generated or otherwise) and to determine its reliability and veracity.

22 With greater usage, and for the reasons given in the preceding paragraph, there will also be greater opportunity for the misuse of Gen AI by feeding harmful or misleading data into, or reusing and disseminating such data from, the Gen AI database. For example, there was a case involving a false report of an accident generated by Gen AI used by a person in China to spread false information.[15]

23 Greater use of Gen AI can also give rise to professional and ethical concerns if the data is used for work purposes, given its capabilities. One recent example was a lawyer using Gen AI-generated submissions in court that cited fictitious cases.[16] Assuming that the judge and the lawyer on the opposing side did not perform due diligence in that case, it could have led to even graver implications, such as bringing disrepute to the institutional trustworthiness and integrity of the legal profession and the justice system, as well as the creation and perpetuation of false precedents and misdirection of future decisions. The same or similar concerns can arise in various other industries where such Gen AI can be useful.

24 Because of the above concerns, it is noteworthy that in its latest version, ChatGPT has acknowledged them and included the following disclaimers:[17]

(a) "Don't share sensitive info. Chat history may be reviewed or used to improve our services."

(b) "Check your facts. While we have safeguards, ChatGPT may give you inaccurate information. It's not intended to give advice."

(c) "ChatGPT can make mistakes. Consider checking important information."

25 ChatGPT also allows the user to "prompt again" for an alternative answer. It will be considered later in Part IV whether these notices and functions are sufficient to deal with the concerns relating to personal data

15 "China Report First Arrest Over Fake News Generated by ChatGPT" *Reuters* (11 May 2023) <https://www.reuters.com/technology/china-reports-first-arrest-over-fake-news-generated-by-chatgpt-2023-05-10/> (accessed 31 May 2024).

16 Sara Merken, "New York Lawyers Sanctioned for Using Fake ChatGPT Cases in Legal Brief" *Reuters* (26 June 2023) <https://www.reuters.com/legal/new-york-lawyers-sanctioned-using-fake-chatgpt-cases-legal-brief-2023-06-22/> (accessed 31 May 2024).

17 See ChatGPT-3.5, available at: <https://chat.openai.com/> (accessed 31 May 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

privacy and false information, and if not, what can and should be done policy wise and by law.

C. Gen AI and personal data

26 The nature of data use for Gen AI models has the broadest data privacy implications. Table 1 below shows when personal data may be involved in the Gen AI process, and the entity or person that constitutes the "data organisation" under the Singapore Personal Data Protection Act 2012[18] ("PDPA") that is responsible for the personal data in its possession or under its control at each stage of the process.[19]

Data controller (action) / **Data "life cycle"**	**Gen AI organisation (generate)**	**User (prompt, reuse)**
Collection and storage	1. Collection and storage of large training datasets as input for models.	4. Saving content from Gen AI output.
Use / processing	2a. Generating content from user prompts. 2b. Improving the model and/or understanding user preferences.	5. Reusing such content in personal or professional capacity.
Disclosure / sharing	3. Presenting content to users.	6. Sharing such content in personal or professional capacity.

Table 1. Points where personal data may be involved

(1) When do Gen AI organisations deal with "personal data"?

27 The data input into Gen AI can include, but is not limited to, personal data. Gen AI services, like many non-Gen AI services, also often collect, store and analyse user interaction to understand user preferences and to thereby improve their services. Hence, training data is continually

18 2020 Rev Ed. Note, this is referred to as the "data controller" in the context of the EU's General Data Protection Regulation: see Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

19 Personal Data Protection Act 2012 (2020 Rev Ed) s 11(2). The reason why the Gen AI organisation is by default the one accountable for data protection obligations (until the content leaves its possession or control) will be given in Part IV.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

renewed and the algorithm is constantly being developed. In the process, personal data about an individual can be *created* by processing the training data through a combination of inductive reasoning and probabilistic predictions.

28 Any point of the collection, use and disclosure ("CUD") process involves personal data. As long as identifiable personal information – *ie*, data that directly identifies an individual or that *can reasonably identify* a person – is involved, then the PDPA will apply. The ease of identification, and when it amounts to personal data, is particularly relevant to the use of Gen AI models as users can actively send prompts to seek information relating to a person as well as receive personal data indirectly from a query not specifically relating to a person. The identifiability of a person can be descriptive (mainly in text form) or through the person's unique features and personal characteristics such as the facial image, voice and gait. For example, Gen AI-generated images that consist of an amalgamation of images, which include the uniquely identifiable features of an individual (often that of a famous person), can constitute personal data if there is sufficient resemblance to that person in question (which is a question of fact).[20] The Singapore Personal Data Protection Commission ("PDPC") has also affirmed in its guidelines that the individual must be able to "be singled out from other individuals … based on one or more characteristics of the data".[21] Another example is using AI to generate audio output (such as an interview, speech, recital or a song) using the AI-generated voice of a person.[22]

29 "Personal data" also has the potential of even broader application and implications in the Gen AI context. For example, Gen AI can be used to make automated *decisions* that can affect a specific person, as

20 Style-based generative adversarial networks can learn facial features from a large dataset of real faces and generate realistic non-existent faces. See Tero Karas, Samuli Laine & Timo Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks" (2021) 43 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 4217.

21 *Advisory Guidelines on Key Concepts in the Personal Data Protection Act* (Personal Data Protection Commission, 24 September 2013) at para 5.9. See also, Warren B Chik & Pang Keep Ying Joey, "The Meaning and Scope of Personal Data Under the Singapore Personal Data Protection Act" (2014) 26 SAcLJ 354 at para 91.

22 It also has to then be determined whether the data is "about" or "refers to" a person. The intentions of the data organisation *vis-à-vis* an individual can be a relevant factor of consideration (*eg*, whether the person is the subject or object and main focus of the image). See *eg*, *Eastweek Publisher Ltd v Privacy Commissioner for Personal Data* [2000] HKCA 442. This approach in Hong Kong is more applicable to the Singapore Personal Data Protection Act 2012 than the General Data Protection Regulation as they both recognise the distinction between personal data protection and the general right to privacy. This contributes to the dynamism of the concept of "personal data" and can narrow its scope somewhat.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

opposed to merely providing information *about* him or her. Gen AI can also produce results that can have an impact on a person, such as its use for evaluative purposes like credit assessment rating by a lending organisation, admission assessment by an educational institution, or an employment decision by a commercial organisation or even the public sector. Because of these, perhaps there should be more restrictions if a Gen AI were used in such a manner that could lead to significant effects on an individual.[23]

(2) What are the rules when dealing with "personal data"?

30 If a Gen AI organisation is dealing with personal data, then the next issue is the origin of the personal data that will determine the legality of the CUD of such data. An important consideration is whether the source from which the personal data was derived was publicly available or from a private source (*eg*, circumventing a firewall or the privacy setting of an online platform).[24] This will depend on the data mining or web scraper tool used to gather the "raw data" for input into the "generator" for processing. The latter may require individuals to be notified and for their permission to be obtained, unless specifically exempted by statute on grounds such as the "publicly available" exception.[25] Furthermore, the nature of the personal data and purpose of its collection can also be relevant since there are statutory exceptions from the consent requirement for the collection and processing of user data to improve and refine the Gen AI.[26]

31 Nevertheless, regardless of the source and purpose, even if exempted from seeking permission or user notification, other PDPA obligations must still be reasonably complied with.[27] For example, there should be user notification (unless reasonably exempted) and the purpose

23 See n 84 below on Art 22 of the General Data Protection Regulation. There is no equivalent provision in the Personal Data Protection Act 2012.

24 In contrast, the retention of chat history for processing to improve the Gen AI service can satisfy the requirements of notification and consent based on the on-site disclaimer (and terms of use and privacy policy) and agreement by conduct (and payment, for premium services).

25 Personal Data Protection Act 2012 (2020 Rev Ed) First Schedule, Pt 2, para 1.

26 Other exceptions that may be relevant and appropriate, depending on the circumstances, are the legitimate interest, business improvement and research exceptions. See the Personal Data Protection Act 2012 (2020 Rev Ed) First Schedule, Pts 3 and 5, and Second Schedule, Pt 2, Divs 2 and 3 (regarding use); and Second Schedule, Pt 3, Div 2 (regarding disclosure).

27 Personal Data Protection Act 2012 (2020 Rev Ed) s 11(1), *ie*, "what a reasonable person would consider appropriate in the circumstances". See in particular, the purpose limitation obligation under s 18(*a*) that restates the same reasonableness threshold requirement.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

of its use must be reasonable and appropriate in the circumstances. Further, the database must be adequately protected according to the PDPA security obligation and data retention is only justifiable as long as the data is needed for any potential query prompted by its user base (or for any other reasonable purpose).[28]

(3) Who is accountable for obligations under PDPA?

32 The third question relates to the allocation of responsibility for PDPA obligations. The answer will depend on whether the Gen AI is considered a "tool" or an independent entity for the purpose of the PDPA, given the fact that it can create content automatically (without involving a "human in the loop", *eg*, human intervention or pre-determination). For reasons that will be elaborated in Part IV, it is most likely to be considered a tool, for which the data is *attributed* to the Gen AI organisation.[29] This is more straightforward for input data since it is programmed to collect data by the organisation's employees,[30] as opposed to the processing and delivery of content that is done automatically by the Gen AI system. It is worth noting that most Gen AI organisations, such as OpenAI, also acknowledge and assume the responsibility of a data organisation (or data controller, in the case of the EU's General Data Protection Regulation) in its terms of use and privacy policy.

33 On the other hand, if a Gen AI model is licenced to another business as a tool for the purpose of managing data that it owns, or at least data that is in its possession or under its control, then that business itself will be accountable as a data organisation under the PDPA and will have to meet the obligations under the Act. If the Gen AI organisation provides support services to that business by processing personal data "on behalf of and for the purposes of" the data organisation "pursuant to a contract which is evidenced or made in writing", then it will be a "data

28 Sections 20, 18, 24 and 25 of the Personal Data Protection Act 2012 respectively. It is questionable whether it is reasonable to require a Gen AI organisation to provide for access to personal data by anyone who asks (s 21 of the Personal Data Protection Act 2012), but the possibility to insist on correction and the importance of accuracy may make it more reasonable than not for the organisation to provide for such a "right", whether manually or technically.

29 Since it does not fulfil the definition of "data intermediary" (s 4(2) of the Personal Data Protection Act 2012) and the Personal Data Protection Act 2012 envisions a responsible entity for any data processing (s 4(3) of the Personal Data Protection Act 2012), it follows that the Gen AI organisation is primarily responsible and accountable for the nature of the generated content. This will be consistent with policy objectives to ensure that government guidelines and laws are adequately complied with, and also that the Gen AI must remain human-centric.

30 Which includes "volunteers" under s 2 of the Personal Data Protection Act 2012.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

intermediary" in this relationship and will only be required to comply with the minimal conditions set out in the PDPA for data intermediaries.[31]

34 In some countries including Singapore, even the public sector is using Gen AI.[32] In a Parliament sitting discussing the development and maintenance of ethical AI standards in the public sector, the Senior Minister of State for Communications and Information stated that:[33]

> Where necessary and useful, we will update our measures to take into account the impact of developments like ChatGPT and GPT-4. For example, the Public Service has introduced guidelines for public officers using similar technologies to draft documents. *These guidelines make clear that public officers are accountable for their work and are responsible for fact-checking and vetting AI-generated content.* The guidelines also aim to safeguard data security by reminding officers not to input sensitive information into these applications. [emphasis added]

Hence, a Gen AI organisation, or a business using Gen AI as a tool, must comply with the law and regulations to address content and privacy concerns, respond to (and co-operate with) the authorities when investigative processes are ongoing and carry out remedial action or comply with regulations as required. These will also be elaborated on in Part IV.

---

31 Personal Data Protection Act 2012 (2020 Rev Ed) s 4(2).

32 Sharanya Pillai, "Singapore Developing Software with ChatGPT-like Features for Public Sector" *Business Times* (18 July 2023) <https://www.businesstimes.com.sg/startups-tech/startups/singapore-developing-software-chatgpt-features-public-sector> (accessed 3 June 2024); Ong Li Min & Jason Grant Allen, "ChatGPT and the Public Service" *GovInsider* (14 July 2023) <https://govinsider.asia/intl-en/article/chatgpt-and-the-public-service> (accessed 3 June 2024). However, it was reported in as early as May 2023 that the Senior Minister of State for Communications and Information affirmed that the official government position recognises that "public officers are accountable for their work and responsible for fact-checking AI-generated content": see Osmond Chia, "Public Officers Can Use ChatGPT and Similar AI, but Must Take Responsibility for Their Work: MCI" *The Straits Times* (23 May 2023) <https://www.straitstimes.com/tech/public-officers-allowed-to-use-chatgpt-and-other-ai-but-must-take-responsibility-for-work-mci> (accessed 3 June 2024). As a "public agency" is excluded from the scope of the Personal Data Protection Act 2012, the allocation and type of responsibility when it comes to the management of personal data is set by other laws and government policy.

33 Singapore Parl Debates; Vol 95, Sitting No 103; [9 May 2023].

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

III. Policy concerns: guidelines, regulations and ethics debate

A. AI governance and regulations

35 Many governments have in recent years begun to introduce policy frameworks, mainly in the form of non-binding guidelines, to steer the development of AI in a manner that recognises and addresses ethical concerns so as to promote "built-in" safeguards for AI in the development process and before its deployment for widespread use. This trend fulfils two complementary goals: (a) to build confidence and trust in AI when it is deployed for societal uses; and (b) to prevent and minimise harm to society while ensuring optimal performance and benefits from its use.[34] At the same time, it will provide a conducive environment for the development of AI that will benefit the economy, which is increasingly technology-focused, and for industries pivoting to the use of AI to achieve greater efficiency and productivity.

36 Prominent examples of such policy frameworks are the EU's Ethics Guidelines for Trustworthy AI ("EGTA"), which was presented by the High-Level Expert Group on AI on 8 April 2019 to highlight the need for lawful, ethical and robust AI systems,[35] and Singapore's Model Artificial Intelligence Governance Framework ("SAIGF") that is in its second version.[36] The former provides the "model" for EU Member States to develop their own guidelines. It also precipitated the development of regulation in the form of the proposed and anticipated Artificial Intelligence Act ("AI Act"), the first proposed AI legislation. The latter is the first set of AI guidelines in Asia and has been adopted by the public and private sector industries in Singapore. In some cases, the general principles in the SAIGF have been adapted into more relevant

34 See Charlotte Stix, "Artificial Intelligence by Any Other Name: A Brief History of the Conceptualization of 'Trustworthy Artificial Intelligence'" (2022) 2(26) Discov Artif Intell 1. See further, Thilo Hagendorff, "The Ethics of AI Ethics: An Evaluation of Guidelines" (2020) 30 *Minds and Machines* 99. For a useful critique on the notion of trust, see Matthias Braun, Hannah Bleher & Patrik Hummel, "A Leap of Faith: Is There a Formula for 'Trustworthy' AI?" (2021) 51(3) *The Hastings Center Report* 17. For one alternative set of principles to build trust (*ie*, beneficence, non-maleficence, autonomy, justice and explicability), see Scott Thiebes, Sebastian Lins & Ali Sunyaev, "Trustworthy Artificial Intelligence" (2021) 31 *Electronic Markets* 447. Usually, these principles do overlap to some extent with fairness, beneficence to humanity (to the individual and society) and openness being commonly touted principles integral to building trustable AI and trust in AI.

35 *Ethics Guidelines for Trustworthy AI*, available at: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (accessed 4 June 2024).

36 *Model Artificial Intelligence Governance Framework* (2nd Ed, 2020), available at: <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf> (accessed 3 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

and specific guidelines for some industries that heavily invest in AI and that are part of Singapore's critical information infrastructure, most notably the financial and healthcare sectors.[37] These are often followed and supplemented by technical and other measures to test and verify that the values and principles set out in the guidelines are met in practice, particularly at the development stage.[38]

B. Policy concerns over AI and lack of global accord

37 At the international level, UNESCO has called for countries to implement its Recommendations on the Ethics of Artificial Intelligence if they have not already done so.[39] This is the first global framework on the matter with particular focus on fighting toxicity (bias, discrimination and stereotyping) and false information (disinformation and misinformation), on the right to privacy and protection of personal data as well as on human and environmental rights concerns. However, these recommendations are not legally binding and do not forge an international consensus for a consistent and harmonised approach to regulating AI under international law, which will be a more effective and concerted approach to addressing AI concerns. The closest effort to achieving a harmonised accord across nations is unsurprisingly made in

---

37 For example, the Monetary Authority of Singapore ("MAS") released a set of principles promoting fairness, ethics, accountability and transparency in the use of AI and data analytics in fintech in 2018, followed by a talent development programme for AI and data analytics, and most recently an open-source toolkit developed by MAS and a consortium of over 30 industry representatives in its Veritas Initiative that began in 2019, see: <https://www.mas.gov.sg/-/media/mas/news/media-releases/2021/veritas-document-1-feat-fairness-principles-assessment-methodology.pdf> (accessed 3 June 2024) and <https://www.mas.gov.sg/schemes-and-initiatives/veritas> (accessed 3 June 2024). The Singapore Ministry of Health produced the AI in Healthcare Guidelines (see <https://www.moh.gov.sg/docs/librariesprovider5/eguides/1-0-artificial-in-healthcare-guidelines-(aihgle)_publishedoct21.pdf> (accessed 3 June 2024)) to guide developers and implementers of AI in healthcare to supplement existing regulations, in particular the regulations governing the development and use of AI medical devices; see: <https://www.moh.gov.sg/licensing-and-regulation/artificial-intelligence-in-healthcare> (accessed 3 June 2024) and <https://www.moh.gov.sg/docs/librariesprovider5/eguides/1-0-artificial-in-healthcare-guidelines-(aihgle)_publishedoct21.pdf> (accessed 3 June 2024).

38 See the "Implementation and Self-Assessment Guide for Organisations (ISAGO)", "Compendium of Use Cases" and the "AI Verify" governance testing framework and software toolkit, available at: <https://www.pdpc.gov.sg/help-and-resources/2020/01/model-ai-governance-framework> (accessed 3 June 2024). Singapore also set up an Advisory Council on the Ethical Use of AI and Data in mid-2018.

39 "Artificial Intelligence: UNESCO Calls on All Governments to Implement Global Ethical Framework Without Delay" *UNESCO* (30 March 2023) <https://www.unesco.org/en/articles/artificial-intelligence-unesco-calls-all-governments-implement-global-ethical-framework-without> (accessed 3 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

the EU with the EGTA and the anticipated AI Act.[40] At the national level, most countries including Singapore have, at least for now, opted out of any form of AI law or regulation, preferring a *laissez-faire* approach and encouraging self-regulation by Gen AI organisations.[41]

38 Despite the proposed AI Act in the EU and the odd proposed AI-centric legislation such as in Canada,[42] there is no clearly discernible trend towards legislating AI generally. Hence, the development of legislation specifically regulating AI in general is at best nascent. As a

40 Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206), available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206> (accessed 3 June 2024). For a general description of the objective and approach of the AI Act, see "EU AI Act: First Regulation on Artificial Intelligence" *European Parliament* (19 December 2023) <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence> (accessed 3 June 2024). The European Parliament reached a provisional agreement with the European Council on the AI Act on 9 December 2023. On 2 February 2024, the Act was unanimously adopted by the Council of EU Ministers, making the goal of the Act becoming law much closer, requiring the European Parliament's approval and publication in the EU's official journal.

41 For example, Singapore's AI governance testing toolkit, AI Verify, is a good example of where the public sector can step in to provide a tool for organisations to help ensure quality and confidence in an AI product or service. See: Josh Lee Kok Thong, "AI Verity: Singapore's AI Governance Testing Initiative Explained" *Future of Privacy Forum* (6 June 2023) <https://fpf.org/blog/ai-verify-singapores-ai-governance-testing-initiative-explained/> (accessed 3 June 2024).

42 Canada is one of the first countries to formally propose AI regulation in the form of the Artificial Intelligence and Data Act ("AIDA"), which was tabled as part of the Digital Charter Implementation Act of 2022 ("Bill C-27"). The proposed AIDA will identify "high-impact AI systems" that will require the developer of these systems to meet certain safety and human rights standards. An AI and Data Commissioner ("AIDC") will be created by the AIDA. The AIDC will be empowered to administer and develop regulations to fulfil the objectives of the Act. See: "Artificial Intelligence and Data Act (AIDA) – Companion Document" *Government of Canada* (13 March 2023) <https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document> (accessed 3 June 2024). It is noteworthy that the AIDC is expected to fulfil many functions similar to that of a privacy or personal data protection commissioner – such as Canada's Privacy Commissioner and Singapore's Commissioner of the PDPC – like education, training and assistance to compliance and enforcement. There will also be criminal law provisions to ensure accountability for risks associated with high-impact AI systems, which is an approach similar to that of the EU's AI Act. This is deliberate as the AIDA is expected to be "inter-operable" with regulations in other jurisdictions, and hence consistent with international approaches and standards for AI regulation. Finally, the Canadian Government has also released the "Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems" in September 2023, available at: <https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems> (accessed 3 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

subcategory of AI, it is even less likely that laws will be developed in the near future to specifically address Gen AI concerns, especially since many of these are common to all categories of AI. Given that the problems and concerns are quite clearly and lucidly identified and generally agreed upon by most policymakers, even if law is developed, it is more likely to be done by sectors through updates to relevant existing legislation.

C. Singapore's accretive approach to governance and regulation of Gen AI

39 In Singapore, the Infocomm Media Development Authority ("IMDA"), a government statutory board under the Ministry of Communications and Information, developed a discussion paper entitled "Generative AI: Implications for Trust and Governance" with Aicadium, a technology company, to examine the risks of Gen AI and propose an accretive and multifaceted approach to resolve problems that may arise, and to manage the aforementioned risks.[43]

40 The paper identified the "six key risks" of Gen AI as: (a) mistakes and "hallucinations"; (b) privacy and confidentiality; (c) disinformation, toxicity and cyber-threats; (d) copyright challenges; (e) embedded bias; and (f) values and alignment. These closely align with the concerns identified by UNESCO, as set out at para 37 above. Singapore's approach is described as "practical, risk-based and accretive", which goal is to achieve a trusted and safe environment for such technology to flourish. Regulation is seen as not of immediate concern until technical tools and standards that can support effective enforcement are developed. Meanwhile, existing government and industry guidelines as well as sectoral laws such as the PDPA, laws against false information and the Copyright Act 2021[44] would suffice to address most of the above issues.[45]

41 Some of the incremental steps beyond the SAIGF to deal with concerns surrounding AI are the AI Verify initiative and industry-led

43 "Generative AI: Implications for Trust and Governance", available at: <https://aiverifyfoundation.sg/downloads/Discussion_Paper.pdf> (accessed 3 June 2024). There is also a *Proposed Advisory Guideline on the Use of Personal Data in AI Recommendation and Systems*, which can have implications for Gen AI if it is used for such purposes (see n 46 below).

44 2020 Rev Ed.

45 See "Fact Sheet – Discussion Paper on Generative AI – Implications for Trust and Governance", available at: <https://www.imda.gov.sg/-/media/imda/files/news-and-events/media-room/media-releases/2023/06/7-jun---ai-annoucements---annex-c.pdf> (accessed 3 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

AI ethical guidelines.[46] In relation to Gen AI, sectoral Gen AI guidelines in the financial industry in the form of the Gen AI Risk Framework have already emerged.[47] The results of these are the "Generative AI Evaluation Sandbox for Trusted AI by AI Verify Foundation and IMDA", the "Cataloguing LLM Evaluations" discussion paper released in October 2023[48] and the "Gen AI Risk Framework" white paper to be published in 2024.

42 Meanwhile, on 16 January 2024, The IMDA, together with the AI Verify Foundation,[49] released a proposed framework specifically for Gen AI, entitled "The Model AI Governance Framework for Generative AI"[50] ("Gen AI Framework"), which will expand on the foundation of the SAIGF and customise it to the issues that are more unique to Gen AI. It will collate feedback from the international community to finalise the Gen AI Framework in mid-2024. The significance is twofold: First, the approach and objective is "systematic and balanced" in order to facilitate Gen AI innovation for socioeconomic benefit while ensuring safety and reducing risk in its use respectively. Second, Singapore is taking the lead in fostering a Gen AI framework that can consequently lead to greater global convergence.[51]

---

46 Information on these can be found in the AI Verify Foundation website, available at: <https://aiverifyfoundation.sg/ai-verify-foundation/> (accessed 3 June 2024). In another relevant development, the IMDA and SG:Digital issued the "Proposed Advisory Guidelines on Use of Personal Data in AI Recommendation and Decision Systems" on 18 July 2023, available at: <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Legislation-and-Guidelines/Public-Consult-on-Proposed-AG-on-Use-of-PD-in-AI-Recommendation-and-Systems-2023-07-18-Draft-Advisory-Guidelines.pdf> (accessed 3 June 2024).

47 "MAS Partners Industry to Develop Generative AI Risk Framework for the Financial Sector" *Monetary Authority of Singapore* (15 November 2023) <https://www.mas.gov.sg/news/media-releases/2023/mas-partners-industry-to-develop-generative-ai-risk-framework-for-the-financial-sector> (accessed 3 June 2024). See also *Emerging Risks and Opportunities of Generative AI for Banks: A Singapore Perspective* (Mindforge, 2024), available at: <https://www.mas.gov.sg/-/media/mas/news/media-releases/2023/executive-summary---emerging-risks-and-opportunities-of-generative-ai-for-banks.pdf> (accessed 3 June 2024).

48 The former is available at <https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2023/generative-ai-evaluation-sandbox> (accessed 3 June 2024) and the latter is available at <https://aiverifyfoundation.sg/downloads/Cataloguing_LLM_Evaluations.pdf> (accessed 3 June 2024).

49 This foundation was launched in 2023 as an initiative involving the major technology companies to help develop AI testing tools for responsible AI use. More information is available at: <https://aiverifyfoundation.sg/> (accessed 3 June 2024).

50 "The Model AI Governance Framework for Generative AI" available at: <https://aiverifyfoundation.sg/downloads/Proposed_MGF_Gen_AI_2024.pdf> (accessed 3 June 2024).

51 Goh Yan Han, "Singapore Seeks International Feedback on New Governance Framework for Generative AI" *The Straits Times* (17 January 2024) <https://>

(cont'd on the next page)

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

43 Table 2 below provides a comparison of the ethical principles and how they are defined in the context of the SAIGF and the Gen AI Frameworks. The principles on accountability and responsibility, accuracy, transparency and explainability, auditability as well as robustness and security will also feature in Part IV in the analysis of the applicability and adequacy of the PDPA and the laws against false information to meet these requirements, to address the aforementioned issues, and to guide the development of these laws (and fill the lacunae by other non-law methods) in areas where they are currently lacking.

AI Governance Framework (2020 version)[52]	**Ethical principle**	**Gen AI Framework (2024 version)**
Ensure that all AI actors are responsible for AI systems respecting AI ethics and principles.	Accountability	Responsibility of Gen AI organisations (and any others involved) to end users by adhering to AI ethics and principles.
Identify false information and mitigate its effects.	Accuracy	Ensuring data quality (including veracity) and resolving training data issues.
Responsibility, redress and record keeping to facilitate self-monitoring or external audits.	Responsibility, accountability and transparency	Trusted development and deployment by enhancing transparency and implementing evaluation and disclosure measures.
Enable third parties to review (monitor/check) and investigate (understand and probe) and take appropriate measures or feedback.	Auditability (ombudsman)	Incident reporting by notification, remediation and improvement measures through the use of an incident management system. External validation, testing and assurance.
AI systems to be safe and protected from abuse.	Robustness and security	Keep track of new threats and update security measures.

www.straitstimes.com/singapore/s-pore-seeks-international-feedback-on-new-governance-framework-for-generative-ai> (accessed 3 June 2024).

52 See Annex A of the *Model Artificial Intelligence Governance Framework* (2nd Ed, 2020), available at: <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf> (accessed 3 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

AI Governance Framework (2020 version)[52]	Ethical principle	Gen AI Framework (2024 version)
Make it understandable to end users.	Explainability	Content provenance (consistent with transparency, identifying source data).
Take measures to avoid and reduce or eliminate discrimination.	Fairness	
Respect human rights.	Human rights alignment	
Take measures to promote benefits to society as a whole.	Human centricity and well-being	AI for public good. Safety and alignment in research and development.
Make AI accessible to all.	Inclusivity	
There is "value-add" in the development and deployment of the AI as opposed to not doing so	Progressiveness (utilitarianism)	
Foresight and long-term lasting benefit	Sustainability	

Table 2. Comparison of the ethical principles in the SAIGF and the Gen AI Framework

IV. Gen AI and personal data concerns

44 The main concerns surrounding personal information relate to Gen AI for public access and use (*ie*, an open system), and not Gen AI used as a tool or service within an organisation (*ie*, a closed system). Hence, unless otherwise stated, the following analysis relates to mainly the former.

A. Overview of stakeholders and Gen AI issues

45 It is important at the outset to identify the main stakeholders to Gen AI, and in particular, their interests and concerns. This will form the basis for the proposals and recommendations for the governance and regulation of Gen AI in the areas of data protection and false information relating to the person:

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

(a) The government as policymaker and regulator is concerned with ensuring that Gen AI serves the larger objectives of providing benefits to society and the economy while avoiding or minimising any negative effects or impact. This requires balancing the interests of different segments of society, in particular individuals and businesses. An example of this approach is the PDPA.[53] As there is heavy use of data and broad dissemination of information through Gen AI, the key elements of the ethical principles in the SAIGF and Gen AI Framework, and more specifically, the data protection principles in the PDPA for personal data, are highly applicable to the issue of data governance and accountability relating to Gen AI organisations. An overview of the trend in Gen AI policy and regulation has already been provided in Part III. As it is likely that the approach to Gen AI issues is sectoral in most countries including Singapore, the applicability of the existing laws and regulations on data protection and false information relating to the person will be examined in this Part.

(b) The Gen AI organisation is an important entrant to the IT or AI industry, and the greater use of Gen AI will have an impact on the economy and many industries as well as on society as a whole. Some sectors may be "disrupted" by the use of Gen AI where it can replace existing services, even as Gen AI can also be used to streamline and improve work processes in others. Meanwhile, there can also be an impact on the workforce through job displacement giving rise to the need for skills upgrading, which is becoming the norm in Singapore whenever new technology is harnessed at the workplace. Also, Gen AI organisations, as innovators and service providers, have an interest in building a good reputation to engender and sustain trust and confidence in its services in order to grow as a viable business. There is also interest in avoiding liability – civil and criminal – and to comply with legal requirements where applicable. One way to do this is for Gen AI organisations to assure the government that their internal practices obviate the need for top-down regulations. Hence, the role and status of such organisations, and the strategy they should take to address social and political concerns (in the form of "best practices" for self-regulation), will be examined in detail in this Part as well.

(c) Last, but not least, the third important stakeholder is the Gen AI user, which can be a natural person or a legal entity. Gen AI may be used in a personal or professional capacity. For non-

53 Personal Data Protection Act 2012 (2020 Rev Ed) s 3.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

commercial personal use, Gen AI can facilitate social and educational (especially research) objectives, which has its own set of concerns. Where professional use is concerned, ethical and professional integrity are issues that have to be addressed. These will be considered when assessing the legal rules and professional standards that should be put in place to maintain the reliability of Gen AI as a tool and organisational integrity within relevant industries.

46 Based on the above analysis and observations, Table 3 below offers an overview of the stakeholders and their concerns over the use of AI generally, and Gen AI specifically.

Social	Data privacy and protection	Sustainability	**Economic**
Human rights and fairness	Workforce displacement	Competitiveness and productivity	Copyright and intellectual property
Socio-political Manipulation	Informational disorder and false information	Trust and confidence in the professional use of Gen AI (ethics)	Gen AI innovation and environment
Political	National defence and security	Security and protection	**Industrial**

Table 3. Taxonomy of Gen AI concerns

47 Many of the concerns in Table 3 do overlap with AI in general, but the shaded portions indicate the more immediate problems that are specific to Gen AI. The taxonomy generally aligns the policy concerns closer to the relevant sector, but the schema is also meant to show that these concerns are not mutually exclusive to any sector. This will explain which issues are better dealt with and by which sector, and the preferable method of governance and controls (*eg*, government regulation, co-regulation or self-regulation). Furthermore, Table 3 is useful to see where personal data-related issues lie in the general scheme of things (see the shaded portions).

48 For Singapore, given the existing comprehensive suite of laws that address the main concerns surrounding Gen AI (including personal data protection and false information laws), a calibrated and incremental approach through the adaptation of existing provisions, where necessary, seems the most appropriate. This is especially as the government's policy is to nurture and provide a conducive environment for Gen AI

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

organisations to set up their businesses, and to innovate in Singapore. The gradated and sectoral approach also seems to be more of the norm and is also the approach taken by major economies like the US and the UK. Moreover, a standalone legislation to regulate AI (and Gen AI) is still the exception rather than the norm according to the above trend analysis.

B. Gen AI as a tool or personality: accountability and responsibility

(1) Role and status of Gen AI organisations

49 New content created by Gen AI, in particular AI language models that respond to user prompts, may be an entirely original creation, or based in whole or in part on information obtained from training data that consists of third-party material. In the context of personal data, Gen AI can create and present information about an individual in a different manner from source material, and even create and share a new narrative about a person.

50 On the one hand, Gen AI for general use currently functions as a form of "pull" technology by responding to individual user prompts, in contrast to mass or social media, which is more aligned to the traditional broadcast media model, that allows for one-to-many dissemination in a short amount of time (*ie*, "push" technology). On the other hand, with the likely increase in its scale of use and further dissemination of its content, the societal impact or outcome can arguably reach the same level so as to warrant making the same or similar legal measures applicable to such Gen AI organisations (*ie*, by holding them accountable and responsible for the actions of the Gen AI). Hence, the primary responsibility for Gen AI-generated information, in the context of this article, should be attributed to the Gen AI organisation for the purpose of the PDPA and false information laws. For the same reasons, the Gen AI organisation should not be immune from legal liability and responsibility under any safe harbour laws, which will be explained further at paras 53–58.

51 There are differences and similarities between Gen AI and "Internet intermediaries" ("IIs") when it comes to the management of personal data. In the functions of a Gen AI, personal data can be involved at all three stages of the CUD process, *ie*, collection, processing (content generation, in the case of Gen AI) and dissemination. In the context of the PDPA, based on the distinction between a "data organisation" and a "data intermediary",[54] the Gen AI is likely to be the former and

54 Personal Data Protection Act 2012 (2020 Rev Ed) s 4(3). An organisation has the same obligation under the Personal Data Protection Act 2012 "in respect of personal

(cont'd on the next page)

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

will have to comply with all its requirements. A data intermediary processes personal data "on behalf of and for the purposes of another organization … pursuant to a contract evidenced or made in writing".[55] It is unlikely that a user can be interpreted as that "other organization" or that the terms of use of a Gen AI constitutes a contract for the purpose of this definition. Also, in their privacy policy for their Gen AI services, Gen AI organisations like OpenAI often already acknowledge that they are data organisations and assume their responsibility as one.[56]

52 In considering whether the existing laws dealing with false information about a person extends to Gen AI organisations, and if so, whether there is any need to specifically adapt these laws (and their measures to deal with them), the objective and impact of Gen AI as well as its functions and nature of operation, as compared to IIs and mass media services, will be considered in greater detail at paras 77–87 below.

(2) Gen AI is not a "technology intermediary" for purpose of safe harbour immunity

53 The Gen AI organisation does not fall under the definition of a "technology intermediary" for the purpose of statutory protections from civil and criminal liability in Singapore law, primarily because it does not merely provide access to or deliver third-party material to end users. Instead, raw input data undergoes processing and is presented as original output content.

54 In the US, there is a generous safe harbour immunity from most criminal and civil liability for "interactive computer services" under a provision of the Communications Decency Act,[57] which states that companies – such as those that operate online forums like Facebook and X (formerly, Twitter) as well as search engines – cannot be considered publishers of the posts that third parties put on their websites and are therefore not liable for such content hosted and shared on their

data processed on its behalf and for its purposes by a data intermediary as if the personal data were processed by the organisation itself".

55 Personal Data Protection Act 2012 (2020 Rev Ed) s 4(2).

56 See the OpenAI General and EU Privacy Policy, available at: <https://openai.com/policies/privacy-policy> (accessed 3 June 2024) and <https://openai.com/policies/eu-privacy-policy> (accessed 3 June 2024) respectively.

57 47 USC § 230(c)(1): "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider." This provision is still active and has survived calls for its removal or reduction of the safe harbour for Internet intermediaries. The US courts, including the Supreme Court of the United States, have consistently interpreted and applied the provision in a manner that reaffirms the extensive protection of IIs from liability for third party content.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

platforms. One of the co-authors of the provision recently said that it does not protect AI chatbots.[58] There are also calls to end the safe harbour immunity for Gen AI when used by a "provider or user of an interactive computer service".[59] These developments indicate that Gen AI organisations themselves are not (and should not) be eligible for such broad protections from liability.

55 In contrast, the narrowly defined general safe harbour provision in the Singapore Electronic Transactions Act 2010[60] ("ETA") is more clearly inapplicable to Gen AI organisations and users of Gen AI tools to generate content. Under s 26(1) of the ETA, a "network service provider" is shielded from:

> … civil and criminal liability under any rule of law [only] *in respect of third-party material* in the form of electronic records to which the network service provider *merely provides access* if such liability is founded on the making, publication, dissemination or distribution of such materials or any statement made in such material; or the infringement of any rights subsisting in or in relation to such material. [emphasis added]

For the PDPA in particular, the provision reiterates that "a network service provider shall not be subject to any liability under the Personal Data Protection Act 2012 *in respect of third-party material* in the form of electronic records to which the network service provider *merely provides access*"[61] [emphasis added].

56 By *creating* content or presenting *modified* content as its output, Gen AI-created data should not be considered "third party material" over which they have "no effective control".[62] A "third-party", "in relation to a network service provider, means a person over whom the provider has

---

58 Meghan McCarty Carino & Rosie Hughes, "Section 230 Co-author Says the Law Doesn't Protect AI Chatbots" *Marketplace Tech* (19 May 2023) <https://www.marketplace.org/shows/marketplace-tech/section-230-co-author-says-the-law-doesnt-protect-ai-chatbots/> (accessed 3 June 2024).

59 Katie Paul, "Bipartisan US Bill Would End Section 230 Immunity for Generative AI" *Reuters* (15 June 2023) <https://www.reuters.com/technology/bipartisan-us-bill-would-end-section-230-immunity-generative-ai-2023-06-14/> (accessed 3 June 2024).

60 2020 Rev Ed.

61 Electronic Transactions Act 2010 (2020 Rev Ed) s 26(2).

62 Similarly, even if a "network service provider" under the Electronic Transactions Act 2010 extends beyond a network connection provider" or a "network access provider" to other IIs including content hosts and information location tools or services as in the case of ss 318 and 319 of the Copyright Act 2021 (2020 Rev Ed), it will not cover Gen AI for the same reason that it creates original or modified material and does not merely transmit third-party material. It should also be noted that Gen AI is unlikely to constitute an "online communication service" under the Broadcasting Act 1994 (2020 Rev Ed), which is defined under s 2A. Moreover, for

(cont'd on the next page)

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

no effective control".[63] A Gen AI is neither a "person" nor one that the Gen AI organisation, which created the Gen AI and continues to develop the application (on the basis of business improvement), does not have "effective control" over when it comes to its content. Gen AI organisations also present output as its own creation and do not attribute or cite sources (unlike, *eg*, search engines and online encyclopaedias).

57 Moreover, Gen AI organisations have the technical knowledge and greater resources to meet the statutory objectives of personal data protection and false information laws than the user (who is concurrently the recipient of Gen AI output and the main subject of these protections) with regard to the data that it processes. Hence, practically, it makes sense to hold them accountable for the output of the Gen AI.

58 In summary, in order to achieve the objectives of the PDPA and false information laws, and for the same reasons that disqualify Gen AI organisations from statutory safe harbours that apply to some technology intermediaries (as provided for in s 26 of the ETA and in the relevant provisions of other written laws), Gen AI is more likely to be considered a *communication tool* rather than a "personality", and its acts should be attributed to the Gen AI organisation, for the purpose of assigning legal liability and responsibility for compliance with regulatory obligations.

(3) Recommendations

59 Having determined that Gen AI is a tool and that the responsibility for its use should primarily lie with the Gen AI organisation or institution using it, what form should this responsibility take beyond compliance with the existing regulations? The following are some proposals on how the government should approach Gen AI from the legal and regulatory standpoint in order to better accommodate Gen AI technology and its assimilation into Singapore's industries, balance the competing interests of stakeholders, and ensure that the ethical principles and broader public policy interests are met:

(a) The government should monitor the impact of Gen AI, and if and when it is determined to be necessary, put in place licensing or accreditation requirements for Gen AI organisations

---

the current purpose of that Act, it only encompasses "social media services": see Broadcasting Act 1994 (2020 Rev Ed) Fourth Schedule.

63 Electronic Transactions Act 2010 (2020 Rev Ed) s 26(4). Also, "provides access" in the context of the immunity provision is defined as "the provision of the necessary technical means by which third-party material may be accessed and includes the automatic and temporary storage of the third-party material for the purpose of providing access".

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

under its content regulations. For example, the Broadcasting Act 1994[64] ("BA") already sets out requirements and conditions for some categories of IIs (like "Internet content providers" and "Internet service providers" that provide "computer online services"[65]) to ensure that the broader objectives of such licensing and regulatory regimes are met. These will also provide assurances to the general public and build trust in the licensed or accredited Gen AI organisations (and in turn, provide the impetus for such organisations to get the permits or endorsements respectively). The licensing and accreditation approach need not be mutually exclusive and the latter approach can involve the participation of industry associations to better lend credence to the standards set for official recognition. An example of an effective certification scheme is the Data Protection Trustmark that measures data organisations under the scope of the PDPA against the obligations of the Act, international benchmarks and industrial best practices. A licensing scheme will better compel good performance due to enforcement measures and penalties for dereliction of duties.

(b) Taking a note from the EU's AI Act and Canada's AIDA, this article argues that there should be laws providing for reporting obligations to fulfil transparency obligations (with protections for confidentiality, especially for the protection of trade secrets and sensitive data). If instituted prior to any licensing scheme, reporting obligations may help the relevant government authority determine whether such a scheme is in fact needed, and voluntary reporting of relevant information will be necessary for any accreditation scheme. If it should be mandated, reporting obligations need not come in the form of a new statute, but can be set out in regulations under existing law, and perhaps incorporated and customised as part of licensing requirements for Gen AI (if that is determined to be the preferred approach, *eg*, a licensable broadcasting service under the BA). Codes of practice can also be issued to guide Gen AI organisations[66] if their internal measures are not adequate or sufficiently consistent.

(c) Similarly, Gen AI organisations can be required, under statutory provisions or licensing requirements, to comply with

64 2020 Rev Ed.

65 See Broadcasting (Class Licence) Notification (GN No S 330/2013). It will have to be determined if including Gen AI organisations under a class licence (in accordance with s 9 of the Broadcasting Act 1994) or an individual licence (under s 8 of the Broadcasting Act 1994) is more appropriate, determined by whether there is a need for more specific conditions.

66 Broadcasting Act 1994 (2020 Rev Ed) s 6.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

existing legal obligations such as the setting up of a proper complaints or notice process to comply with existing laws. For example, a procedure that responds to private complaints or government directions to take down, remove, block or deny access to any form of unlawful content.[67]

(d) Especially to combat the dissemination of (and belief in) false information, some changes can be made to improve on the current best practices. These can include requiring citation of sources of information for all Gen AI open system platforms in the presentation of output information, so that the user-recipient has the opportunity to check assertions against the source, verify sources from which data is drawn according to trustworthiness of the source measured by a well-respected reputation index (*eg*, a source that has a robust fact-checking mechanism, preferably by a neutral third party), and there can be oversight over the Gen AI by a neutral third party such as an appropriate government agency (like the POFMA Office[68]). This will also promote discernment in Gen AI users (as the primary recipients of Gen AI information), thereby indirectly reducing the damage that false information can cause (to society or a data subject).

C. Gen AI and "personal data": identifiability, inferences and impact

(1) Direct and indirect identification of the person and "motivated investigator test"

60 An important element of "personal data" is the *identifiability* of an individual that, based on the definition in the PDPA, is set to a reasonableness standard. This is based on: First, the objective of the Act to maintain a fair balance of interests between individuals and businesses;[69] and second, the "reasonableness standard" which appears not only in the general compliance provision,[70] but also in the purpose

67 On condition that it falls under the free speech exception in Art 14(2)(1) of the Constitution of the Republic of Singapore (2020 Rev Ed), as is required of all content regulations. Parliament can by law impose "such restrictions as it considers necessary or expedient in the interest of the security of Singapore or any part thereof, friendly relations with other countries, public order or morality and restrictions designed to protect the privileges of Parliament or to provide against contempt of court, defamation or incitement to any offence".

68 The Office is responsible for the administration of the Protection from Online Falsehoods and Manipulation Act 2019 (2020 Rev Ed).

69 Personal Data Protection Act 2012 (2020 Rev Ed) s 3.

70 Personal Data Protection Act 2012 (2020 Rev Ed) s 11.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

provision.[71] Hence, a purposive interpretation of the phrase "who can be identified" in the definition of "personal data" should include both direct and indirect identification of an individual *with reasonable diligence*, to an extent that is *practicable* and *appropriate*, taking into consideration factors such as the state and sophistication of Gen AI technology, the intention and interest of the user, and the user's proficiency in the use of Gen AI to identify someone from the information available to them.

61 This is especially so since personal data can be presented by Gen AI in a way that also permits indirect identification of a person through reasonable deduction and inference such as from innuendo or insinuation. Personal data can and should clearly include inferred personal information because the impact in the person is indistinguishable from directly identifying data.[72] Inferences made by machine can be the basis for indirectly identifying a person, and it can also create a narrative about a person and contribute to the "profiling" of a data subject for and by the user-recipient.[73] In the case of a Gen AI-generated image or voice, the similarity of the image and voice to a person, especially if that person is famous or has a very unique visage or voice, the greater the likelihood that the actions or words can be attributed to him or her.

62 A useful approach, which will be especially relevant to Gen AI that can increase incidents of indirect identification, is to adopt and use a test similar to the "motivated intruder" test, which was put forth by the PDPC as a general test to assess the risk of reidentification of an individual from ostensibly "anonymised" data.

---

71 Personal Data Protection Act 2012 (2020 Rev Ed) s 3.

72 This will make it in line with the recent European Court of Justice ("ECJ") decision in *OT v Vyriausioji tarnybinės etikos komisija* Case C-184/20, ECLI:EU:C:2022:601, where sexual orientation can be (reasonably) inferred or derived from name-related data on the spouse, cohabitant or partner by deduction and mental combination, available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62020CJ0184> (accessed 3 June 2024). The ECJ's decision thus extends the scope of application of Art 9 of the General Data Protection Regulation on the processing of special categories of personal data, which are subject to more stringent rules, to "indirectly sensitive data".

73 It is noteworthy that the PDPC is in the process of gathering feedback on "how the PDPA applies to the collection and use of personal data to develop and deploy AI systems that embed machine learning (ML) models ('AI Systems') used to make decisions, recommendations or predictions": see "Public Consultation for the Proposed Advisory Guidelines on Use of Personal Data in AI Recommendation and Decision Systems" *PDPC* (18 July 2023) <https://www.pdpc.gov.sg/guidelines-and-consultation/2023/07/public-consultation-for-the-proposed-advisory-guidelines-on-use-of-personal-data-in-ai-recommendation-and-decision-systems> (accessed 3 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

63 According to the Advisory Guidelines on the Personal Data Protection Act (PDPA) for Selected Topics:[74]

> The 'motivated intruder' test considers whether individuals can be re-identified from anonymised data by someone who is motivated, reasonably competent, has access to standard resources (e.g. the Internet and published information such as public directories), and employs standard investigative techniques (e.g. making enquiries of people who may have additional knowledge of the identity of the data subject).
>
> As anonymisation is a package of risk control measures tailored to the purpose of disclosure and the recipient, the 'motivated intruder' test has to take into consideration the known and possible motivation and resources of the intended recipient organisation.

64 Whereas the motivated intruder test is used to assess the robustness of anonymisation techniques to *protect* personal data (and minimise harm to data subjects),[75] it can also be adapted and used to achieve the same objective by determining whether the information *constitutes* personal data in the first place so as to ensure that the rights of the data subject, including security rights, arise in the first place.

65 In order to distinguish the two, it is proposed that a "motivated investigator test" should be used to determine identifiability of a person. The factors that will be relevant can be similar and include the profile of the data subject (*eg*, whether he or she is a public figure and has unique traits), the profile and intention of the user *searching for* content (such as tech-savviness and whether he or she is actively seeking, or is likely to seek, information on the data subject), the motivation of the user when using the Gen AI to create content (*eg*, for producing deepfakes or mimicry relating to a data subject), the capabilities of the Gen AI used, the context surrounding the identifiable data, and the accuracy and sensitivity of the information (for which greater weight should be given to the interests of the individual, since privacy impact assessment is increasingly a feature for the reasonableness assessment under the PDPA).

66 As noted, such information can be false or can include sensitive information, which can have a greater impact on the individual, especially at the output stage. In such a case, the definitional requirement that the information is "about" or "relates to" a person will be strengthened as well, if it is accepted that it includes the possible impact or effects (especially

74 (Personal Data Protection Commission Singapore, 23 May 2014) at paras 3.37–3.38.
75 *Advisory Guidelines on the Personal Data Protection Act (PDPA) for Selected Topics* (Personal Data Protection Commission Singapore, 23 May 2014) at para 3.14.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

negative or potentially adverse) on the person.[76] Such an interpretation also reinforces the ethical principles for the protection of human rights, serves to ensure human-centric AI or Gen AI processes, and will help protect humans against discrimination and bias in AI or Gen AI systems (see Table 2 at para 43 above).

67 In summary, this proposed broader interpretation to the definition of "personal data" will give greater rights and control to the data subject over his or her rights under the PDPA and related laws, including laws for the protection of personal information from misuse and the protection of the person from false light information.[77] This is because this broader interpretation of "personal data" can also (and is more likely to) be adopted for the provisions in other legislation in relation to "personal data", "personal information" and "identity information".

(2) Rights of access and correction (and obligation to ensure accuracy) under PDPA

68 Embracing a broader interpretation of "personal data" gives an individual access to existing mechanisms to deal with inaccurate information about him or her.[78] It allows for the correction (including, more practicably, the removal) of information about the person that is false or if the purpose of the CUD process is not reasonable (*eg*, creating deepfakes or AI-generated images or videos of a person as a prank but in

76 See: Article 29 Data Protection Working Party, *Opinion 4/2007 on the Concept of Personal Data* (01248/07/EN WP136, 20 June 2007) at p 10, available at: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf> (accessed 3 June 2023). This Opinion is not binding but is often referred to, even after the Working Party was replaced by the European Data Protection Board, which has not issued a new set of guidelines to supersede the Opinion. It remains a useful basis to determine the "personal data" and is consistent with the generous interpretation given by the European Court of Justice. See also, Nadezhda Purtova, "The Law of Everything: Broad Concept of Personal Data and Future of EU Data Protection Law" (2018) 10(1) *Law, Innovation and Technology* 40.

77 For example, criminal liability for the misuse of private data without consent or collected unlawfully is provided in s 416A of the Penal Code 1871 (2020 Rev Ed) (under the heading "Illegally obtained personal information") and s 9 of the Computer Misuse Act 1993 (2020 Rev Ed) (under the header "Supplying, etc., personal information obtained in contravention of certain provisions") respectively. Common law action is also available for the control of personal information to a narrower extent, such the law on defamation and malicious falsehood, breach of confidence, and harassment (including doxing).

78 Personal Data Protection Act 2012 (2020 Rev Ed) s 21. The organisation must do so "as soon as reasonably possible" unless a statutory exception applies.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

a way that is harmful to a person's reputation and mental well-being is arguably unreasonable use of personal data under the PDPA).[79]

69 In order to even discover if personal data, including sensitive information or false information, can be generated by a Gen AI, access is required. However, there is no need for special arrangements as access can be made through the use of the Gen AI service to search for information about oneself (the colloquial term for this act is "egosurfing"). This is the more practicable approach to the access obligation than requiring Gen AI organisations to give access to their information residing in the training, and especially since personal data can be created by the Gen AI that is not necessarily reflective of the source data stored in its system. If this approach is taken, the access obligation can be easily satisfied by Gen AI organisations.

70 In relation to untruthful data about the person, the correction obligation in the PDPA should be interpreted broadly to include the removal or deletion of false information.[80] This provides a useful recourse for individuals as they can rely on the resources of the PDPC instead of taking a civil action against a perpetrator, which can be prohibitive due to the cost and time involved.[81]

71 The Gen AI organisation will also be required to develop their model or program to detect and remove false information to comply with the accuracy obligation under the PDPA.[82] This is already a condition in some ethical frameworks and laws. For example, according to the proposed EU AI Act, and as part of the transparency requirement, Gen AI must

79 Personal Data Protection Act 2012 (2020 Rev Ed) s 18. This is especially so when the impact assessment is becoming an increasingly important measure in determining reasonableness of an action or inaction (exception), such as in determining whether the purpose limitation is breached or whether an exception to consent such as the "legitimate interest" basis is made out.

80 Personal Data Protection Act 2012 (2020 Rev Ed) s 22. An individual may request an organisation to correct an error or omission in the personal data about the individual that is in the possession or under the control of the organisation" and it must do so "as soon as practicable": see ss 22(1) and 22(2)(*a*) of the Personal Data Protection Act 2012. The organisation can only not comply if it is satisfied on reasonable grounds that it should not be done or if a statutory exception applies.

81 The PDPC can take relevant measures such as issuing penalties and directions, but it may be noted that the data subject also, based on the determination of non-compliance with the PDPA obligations and if it can be proven that "loss or damage" is suffered, has a right of action for relief in civil proceedings in court (see s 48O of the Personal Data Protection Act 2012).

82 Personal Data Protection Act 2012 (2020 Rev Ed) s 23. The standard is one of "reasonable effort" and if it is likely to be used to make a decision that can affect the person concerned or is likely to be disclosed to another organisation: see ss 23(*a*) and 23(*b*) of the Personal Data Protection Act 2012.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

alert Gen AI users that the content is AI-generated and design the model to prevent it from generating illegal content. The second requirement can be aligned with the statutory obligation to take reasonable measures to maintain the accuracy of personal data or be subject to sanctions under the PDPA or other laws such as the PHA (which also provides remedies in response to false information about a person, see para 84 below).

72 In contrast, and as a counterweight to the above obligations, it may be noted that there are several broad exceptions that exempt the Gen AI organisation from having to seek permission for the CUD of personal data in the first place, which relinquishes it from an otherwise onerous consent (and notification) obligation. These include the CUD of personal data without consent for the purpose of business improvement and research.[83]

(3) Further suggestions

73 It was mentioned earlier at para 45(c) that the concerns are different whether Gen AI is used for non-commercial personal use or for commercial and professional use.

74 In so far as education, awareness and training are concerned for the Gen AI user (and the data subject potentially identified in Gen AI output), the Commissioner of the PDPC and other relevant government agencies (like the POFMA Office), are in the best position to expand on their role to fulfil these objectives. Such instruction and guidance can advise the user on the responsible use of Gen AI, and the subject on their rights *vis-à-vis* output, as it relates to the latter's personal data. On the training side, understanding of Gen AI and its responsible use can also be built into an accreditation or licensing regime as recommended at para 59 above.

75 Additionally, for Gen AI use by the private sector, it will be more appropriate for professional institutions within specific industries or sectors to set the rules and standards when it comes to the use of Gen AI at work. For example, the Law Society of Singapore can do so for the legal sector, the Singapore Medical Council, Singapore Medical

83 For example, in relation to training data, the "legitimate interest" and "research" exceptions (see s 17 and the First and Second Schedules of the Personal Data Protection Act 2012) can be applicable, and in relation to user data, the deemed consent and "business improvement" exception can apply (see ss 15 and 17, and the First and Second Schedules of the Personal Data Protection Act 2012 respectively). Once there is deemed consent or a statutory exemption applies, notification is also dispensed with generally according to s 20(3) of the Personal Data Protection Act 2012.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

Association and/or the Centre for Medical Ethics and Professionalism for the medical sector, and the professional bodies of other sectors like finance and insurance for their respective industries. Industry standards can be the basis for an accreditation scheme. To maintain standards, disciplinary proceedings can be activated to investigate and sanction a person or entity for non-compliance (to the extent that the institution is legally empowered to do so). These institutions could, likewise, provide training and education on the legitimate and appropriate uses of Gen AI, define the boundaries for the use of Gen AI, and provide guidelines for its use.

76 Related to the above suggestion, it should be considered if automated decision making that has an impact on a person, such as the use of personal data obtained from Gen AI to automatically evaluate a person for the purpose of loans (*eg*, credit rating) and admission to an institution or employment, should be subject to stricter *restrictions* such as explicit consent and independent verification.[84] In fact, it should first be considered by professional institutions if rules should be put in place to *prohibit* the use of Gen AI output by a decision-maker to even make such decisions, especially if it is necessary to protect the rights of the data subject from possible discrimination and arbitrariness in outcome (which are against the principle of using AI for the well-being of humans and the requirement of human centricity).[85] Consideration should also be given as to when it may be appropriate for a "right of erasure" (also more popularly known as the "right to be forgotten") to be incorporated into the PDPA, or at least in industrial data governance practices, as greater technological incursions into personal data privacy are observed.[86]

---

84 For example, Art 22 of the General Data Protection Regulation (entitled "Automated individual decision-making, including profiling"). It states that "a data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her" except for three narrow exceptions including explicit consent.

85 In *Policymaking in the Pause* (Future of Life Institute, 12 April 2023) at p 13, the Future of Life Institute recommended:

> … the establishment of laws and industry standards … [requiring] the fulfilment of 'duty of loyalty' and 'duty of care' when AI is used in the place of or in assistance to a human fiduciary … Laws and standards towards this end should require that if an AI system is to contribute to the decision-making of a fiduciary, the fiduciary must be able to demonstrate beyond a reasonable doubt that the AI system will observe duties of loyalty and care comparable to their human counterparts. Otherwise, any breach of these fiduciary responsibilities should be attributed to the human fiduciary employing the AI system. [emphasis in original omitted]

86 For example, Art 17 of the General Data Protection Regulation (entitled "Right to erasure ('right to be forgotten')"). It provides several bases that give a data subject the right to require a data controller to erase his or her personal data concerning without

(cont'd on the next page)

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

D. Gen AI and informational disorder: veracity and transparency

(1) False information or "hallucinations" and its impact on the person

77 It is possible that the falsehood in Gen AI output is derived from the training data, but it is also possible for the Gen AI to fabricate stories and construct falsehoods through its algorithm. Societal concerns arise from the dissemination of false information that are or may be harmful or that may have a negative impact on the person, institution (whether private or public) and/or society as a whole. Specifically in relation to the person, false information can affect their reputation and mental well-being, and can also give rise to severe forms of harassment such as doxing. Institutions can also face negative repercussions such as reputational harm and economic losses and in the case of public institutions, loss of confidence as well as political misfortune.

78 Examples have emerged from around the world of these effects of false information from Gen AI output, even as the use of Gen AI gains traction and popularity. For example, a mayor of a Victorian country town in Australia threatened action for defamation against OpenAI for falsely stating that he was imprisoned for bribery when in fact he was the whistleblower for a bribery case.[87] A similar reported defamation case also arose in the US.[88] In yet another case, an academic was accused of sexual harassment based on a news article that did not exist.[89]

79 Another more indirect way that Gen AI can be abused is when it is used as a tool to generate false information for the purpose of downstream sharing or dissemination. There was such an incident in China relating to

undue delay. This is not an absolute right and is subject to countervailing interests. Hence, if such a right is added to the Personal Data Protection Act 2012, business concerns (as well as public interest) can be made exceptions to the right.

87 Nick Bonyhady, "Australian Whistleblower to Test Whether ChatGPT Can Be Sued for Lying" *The Sydney Herald* (5 April 2023) <https://www.smh.com.au/technology/australian-whistleblower-to-test-whether-chatgpt-can-be-sued-for-lying-20230405-p5cy9b.html> (accessed 3 June 2023).

88 Isaiah Poritz, "OpenAI Fails to Escape First Defamation Suit from Radio Host" *Bloomberg Law* (17 January 2024) <https://news.bloomberglaw.com/ip-law/openai-fails-to-escape-first-defamation-suit-from-radio-host> (accessed 3 June 2024).

89 Pranav Dixit, "US Law Professor Claims ChatGPT Falsely Accused Him of Sexual Assault, Says 'Cited Article Was Never Written'" *Business Today* (8 April 2023) <https://www.businesstoday.in/technology/news/story/openai-chatgpt-falsely-accuses-us-law-professor-of-sexual-harassment-376630-2023-04-08> (accessed 3 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

fake news of a train crash that did not occur.[90] In this case, it was the user that was abusing the Gen AI to create and disseminate false information. Similarly, using AI to create deepfakes like false images (still or moving, with or without audio) of a person can give rise to a cause of action under existing laws such as those that protect the reputation (*eg*, the torts of defamation or malicious falsehood) and intellectual property rights (*eg*, image or personality rights, only available in in certain jurisdictions) as well as laws against impersonation (*eg*, cheating by personation).[91]

(2) Relevance of Protection from Online Falsehoods and Manipulation Act 2019 and usefulness of Protection from Harassment Act 2014

80 In relation to false information relating to a person, it has been explained earlier how the rights of access and to correction of personal data under the PDPA can already be useful to the individual and provide the necessary recourse and tools for such information to be corrected or removed.

81 Furthermore, for individuals in political office or public servants, the Protection from Online Falsehoods and Manipulation Act 2019[92] ("POFMA") also provides some useful measures, if the false statement of fact could incidentally affect the outcome of political elections or cause a diminution of public confidence in the performance of a public duty or function. This process is expedited as it only requires the relevant Minister to make the assessment that it is "necessary or expedient" to take measures to protect against the above outcome. Such measures consist of the issuance of a direction to the "communicator" (under Pt 3 of the Act), or to IIs and "providers of mass media services" (under Pt 4 of the Act), of false statements of fact.[93]

82 A Gen AI organisation is not an II within the scope and definition of an "internet intermediary" in the POFMA for Pt 4 to apply. For the purpose of the Act, an II refers to a person or entity that provides any Internet intermediary service, which is further defined as one that facilitates or allows end users to find and access *third-party material* or that transmits such materials to end users.[94] The implication is that it is

90 "China Arrests ChatGPT User Who Faked Deadly Train Crash Story" *The Business Times* (9 May 2023) <https://www.businesstimes.com.sg/international/china-arrests-chatgpt-user-who-faked-deadly-train-crash-story> (accessed 3 June 2024).

91 Penal Code 1871 (2020 Rev Ed) s 416.

92 2020 Rev Ed.

93 Protection from Online Falsehoods and Manipulation Act 2019 (2020 Rev Ed) ss 10 and 20.

94 Protection from Online Falsehoods and Manipulation Act 2019 (2020 Rev Ed) s 2.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

a primarily a conduit and does not deliver to end users original content, whether created or modified.[95]

83 Instead, Pt 3 of the Act can cover Gen AI if it communicates an alleged false statement of fact, and there is only liability for non-compliance with the directions. Hence, there is no need to amend the Act to cover Gen AI organisations. The existing stop communication direction may be the most suitable measure to use, whereas the correction direction may not be as useful, given how and when Gen AI currently presents output data (*ie*, usually directly to an individual user-recipient in response to his or her prompting).

84 Similarly, the remedies provided to deal with false information to an individual or entity under Pt 3, Div 2 of the Protection from Harassment Act 2014[96] ("PHA") may be useful against falsehoods generated by Gen AI. However, as it requires bringing an action in court, in reality it may not be the easiest recourse for individuals (and even institutions) to take. On the other hand (and as noted at paras 68–72 above), for individuals, but only in relation to personal data, the PDPA already gives individuals the right to demand reasonable access and correction of inaccurate personal data in the possession and control of the Gen AI organisation, and it also gives the person access to the PDPC complaints process whereby the PDPC will use its resources to investigate and resolve the matter on his or her behalf.

(3) Recommendations

85 In response to the hazards of anthropomorphising AI technology that can mimic human writing style, speech and actions very well, some countries have already enacted laws that require disclosure of Gen AI-generated content to ensure transparency in order to promote discernment on the part of information recipients and to enhance the protection of data subjects. For example, Art 52 of the draft EU AI Act requires AI systems to be designed in such a way that users are informed that they are interacting with an AI. Similarly, the State of California has enacted a similar "bot disclosure law" in 2019.[97] At the international level, the Member States of the UN have adopted UNESCO's Recommendation on the Ethics of Artificial Intelligence, which requires Member States to ensure that AI users "can easily identify whether they are interacting

95 The examples provided also support this: social networking, search engine, content aggregation, internet-based messaging and video-sharing services. See s 2 of the Protection from Online Falsehoods and Manipulation Act 2019.

96 2020 Rev Ed.

97 Senate Bill No 1001, Chapter 892, available at: <https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201720180SB1001> (accessed 3 June 2024).

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

with a living being, or with an AI system imitating human or animal characteristics, and can effectively refuse such interaction and request human intervention".[98] The Future of Life Institute also "recommend[ed] the expansion of 'bot-or-not' laws that require disclosure when a person is interacting with a chatbot. These laws help prevent users from being deceived or manipulated by AI systems impersonating humans and facilitate contextualising the source of the information."[99]

86 As an example of good practice, ChatGPT already provides notices and disclaimers.[100] Perhaps such notices should also be made mandatory (*eg*, as a licencing requirement) or at least promoted as a form of best practice (perhaps for accreditation status).

87 More effective tools, both technical and regulatory, are needed to assist the public in evaluating the authenticity and veracity of the content that they consume. On top of measures for identifying and managing AI-generated content, additional ways to further the objectives of minimising the negative effects of false information, as well as to burst the information filter bubble, is to recommend or require the provision of citation and attribution (to identify the source), and to minimise opinion-based output (and opinions disguised as facts). To deal with challenges posed by veracity concerns, it is also worth reiterating that rules requiring reporting obligations on its operations and the data it uses, and mandating the setting up of an accessible complaints mechanism to comply with the existing laws, can be useful. Especially relating to combating the belief in and dissemination of false information, technical and non-technical methods to verify the authenticity and accuracy of content should be considered (especially by a neutral third party).

V. Conclusion

88 Some adjustments to the law are needed to accommodate Gen AI and incorporate its use in social and professional settings, to maximise its benefits while minimising adverse effects. But there is currently no need for an extensive or comprehensive law reform to tackle the issues that arise from Gen AI, since it is by and large an extension of AI and the main issues have been around for some time (and are already covered by existing laws to an extent).

98 "Recommendation on the Ethics of Artificial Intelligence" (UNESCO, 23 November 2021) at p 37, para 127 <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence> (accessed 3 June 2024).

99 *Policymaking in the Pause* (Future of Life Institute, 12 April 2023) at p 12.

100 See para 24 and n 17 above.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.

89 The regulatory authority, in particular the IMDA, and Gen AI organisations should build a symbiotic co-operative relationship, since Gen AI can be very beneficial to society, the economy and relevant industries. The former can set rules and guidelines to meet policy interests and the fundamental ethical principles, and to balance the interests of all the stakeholders while providing education and training for users and the data organisations respectively, in order to better address the issues surrounding Gen AI. The government can provide a regulatory "sandbox" for the Gen AI industry to self-regulate but monitor how Gen AI organisations deal with the concerns arising from the use of their Gen AI in order to determine if their measures are adequate. The possibility of a licencing and/or accreditation scheme should be kept in consideration, and activated when it is determined to be necessary to advance national interests (and to better maintain the balance of interests of all parties). Hence, it behoves Gen AI organisations to put in place "best practices" to address current concerns, if they want to avoid more top-down regulations and mandatory compliance processes, and for the industry as a whole to co-ordinate their response for a more consistent and concerted approach.

90 Meanwhile, and at the international level, a multilateral harmonised and consistent approach should be developed. As it has the broadest membership, perhaps a UN agency can be assigned such a role. In the meantime, the guidelines and recommendations by international organisations like UNESCO serve to encourage greater dialogue on matters of concern surrounding the use of Gen AI and the most appropriate solutions to these issues.

© 2024 Contributor(s) and Singapore Academy of Law.
No part of this document may be reproduced without permission from the copyright holders.