

**Guidelines on the Use of Generative Artificial Intelligence
for Judges and Judicial Officers and Support Staff
of the Hong Kong Judiciary**

PURPOSE

These guidelines seek to set out general rules and guiding principles for Judges and Judicial Officers (“JJOs”) and support staff of the Judiciary on the use of generative artificial intelligence (“AI”) in performing judicial and administrative duties.

GENERAL PRINCIPLES ON MAKING USE OF TECHNOLOGY IN COURT

2. The primary role of the Judiciary is to administer justice and adjudicate cases in accordance with the law, upholding the rule of law and safeguarding the rights of the individual. Any use of AI (including generative AI) should follow the principles on the use of technology in court operations. Specifically, technology serves to support the Judiciary in performing its role and functions more effectively and efficiently, without compromising the principles of judicial independence, impartiality, and accountability. The use of technology must not be allowed to undermine the dignity and standing of the judicial office or the public trust and confidence in the administration of justice. Before using any AI, it is important for the Court to understand and assess the AI’s capabilities and limitations, as well as address the potential risks involved. In short, any use of AI by JJOs and staff must be consistent with the Judiciary’s overarching obligation to protect the integrity of the administration of justice.

GENERATIVE AI

3. AI generally refers to computer systems which can perform tasks normally requiring human intelligence. Generative AI is a form of AI that generates new content, such as texts, images, or other media, based on pre-training datasets. Generative AI chatbot is a computer programme that simulates an online human conversation using generative AI. Some of the other common terms relating to AI is set out in the *Annex* to these Guidelines.

4. While generative AI has the potential to be a powerful tool for creativity and innovation, we should take note of the potential concerns and challenges when adopting the technology. In particular, generative AI should not be used in any way that may result in a contravention of prevailing laws, regulations or court orders.

GUIDING PRINCIPLES ON RESPONSIBLE USE OF GENERATIVE AI

5. Subject to the general rules and guiding principles set out in these guidelines, JJOs and the Judiciary's support staff may make prudent and responsible use of generative AI in the course of their work where appropriate.

(A) No Delegation of Judicial Functions

6. JJOs should ensure that all judicial decisions continue to be independently and personally made by themselves, and should not under any circumstances allow generative AI to take over performance of their judicial functions. In other words, the Court must ensure that any use of generative AI does not usurp or encroach upon its judicial functions but merely supports and facilitates their performance.

(B) Understand the Limitations of Generative AI; Check to Ensure Accuracy and Accountability

7. Generative AI is a fast developing technology. There will be an increasing range of products available. It is important for a user to understand the features and limitations of the particular model being used. For example, currently many generative AI chatbots are based on large language models and generate new texts (and images or other media) using complex algorithms based on the prompts they receive and the data they have been trained upon. The output generated is what the model predicts to be the most likely combination of words and data, based on the documents and data held as source information, although there can be an element of randomness in the way a model responds. The quality of the output depends on how the generative AI chatbots have been trained, the reliability of the training data and the quality of the prompts entered. They may not necessarily provide answers from authoritative databases. Beware that even with the best prompts, the output can be inaccurate, incomplete, misleading or biased. For illustration, some AI tools may (the following is not exhaustive) –

- (a) make up fictitious cases, citations or quotes, or refer to legislation, articles or legal texts that do not exist – a risk stemming from the fact that large language models can “hallucinate”;
- (b) provide incorrect or misleading information regarding the law or how it might apply;
- (c) make factual errors; and
- (d) confirm that the information is accurate if asked, even when it is not.

8. JJOs and support staff must be mindful of the capabilities and limitations of the generative AI used and check and verify any information obtained to ensure its accuracy and reliability before using or relying upon it in their work. Using information generated without appropriate checking and verification risks causing injustice and damaging public confidence in the Judiciary.

(C) Maintain Information Security; Uphold Confidentiality and Privacy

9. In order to maintain information security, JJOs and support staff should only use information technology (“IT”) devices provided by the Judiciary (rather than personal devices which may not have proper safeguard on information security) to access the generative AI tools. Do not connect IT devices provided by the Judiciary to untrusted networks (including WiFi), particularly those available at public places. Auto-connection or log-in to unknown WiFi should not be enabled. Use your Judiciary e-mail address in your work to maintain information security.

10. Some generative AI chatbots retain the information you input and use it to respond to queries from other users. Unless you are using closed-end generative AI, it should be assumed anything you input can become publicly known. JJOs and support staff should not enter any information which is private, confidential or sensitive into open or public generative AI chatbots. Make sure that your input is adequately generalised and anonymised. Disable the chat history function in the chatbots if this option is available. Take note that some AI platforms may request various permissions which give them access to information on the IT devices used to access the platforms. All such permissions should be refused.

11. JJOs and support staff should ensure compliance with the requirements under the Personal Data (Privacy) Ordinance (Cap. 486), including the six Data Protection Principles (“DPPs”)¹ in Schedule 1 thereto, when handling personal data in the course of using generative AI. In the event of any suspected breach of information security or privacy following the use of generative AI for judicial or administrative duties, the JJO concerned should report the incident to his/her Court Leader as soon as possible, and support staff concerned should report the incident to his/her supervisor as soon as possible, who will then report to the relevant Division Head. For suspected breach of personal data privacy, the Judiciary Data Protection Officer (currently Assistant Judiciary Administrator (Quality and Information Technology)) should also be informed.

(D) Guard against Infringement of Copyright and Intellectual Property Laws

12. JJOs and support staff should avoid using generative AI in any way which may infringe copyright and contravene intellectual property law. For instance, uploading any published materials covered by intellectual property to a generative AI chatbot to obtain a summary or analysis could breach the author’s copyright. Copyright issues may also arise from outputs that are extracted from an original work. It is the user’s responsibility to ensure compliance with copyright and other intellectual property laws² when using generative AI.

¹ The six DPPs cover the entire life cycle of the handling of personal data from collection to destruction, including: (1) purpose and manner of collection; (2) accuracy and duration of retention; (3) use of data; (4) data security; (5) openness and transparency; and (6) access and correction. For details of these DPPs, please refer to Appendix A of “AI: Model Personal Data Protection Framework” recently published by the Office of the Privacy Commissioner for Personal Data (the website link is provided in footnote 3(a) below).

² These include, in Hong Kong, the Copyright Ordinance (Cap. 528), Prevention of Copyright Piracy Ordinance (Cap. 544), Trade Marks Ordinance (Cap. 559), Trade Descriptions Ordinance (Cap. 362), Patents Ordinance (Cap. 514) and Registered Designs Ordinance (Cap. 522).

(E) Be Aware of Bias

13. We should be aware that generative AI chatbots generate responses based on the dataset they are trained on. The responses generated will inevitably reflect any biases (including cultural or ethical biases) or geographical focus or misinformation in the training data. JJOs and support staff should take note of this and make necessary corrections before using or relying upon the information generated.

(F) Take Responsibility

14. JJOs and support staff should bear in mind that ultimately they are personally responsible for any materials produced in their name even where information derived from generative AI has been used.

(G) Be Aware of the Use of Generative AI by Court Users

15. We should be aware of the possibility that court users have used generative AI in the preparation of litigation documents or materials. While lawyers have a professional obligation to ensure that any material they present to the Court (however generated) is accurate and appropriate, JJOs should, in appropriate cases, remind individual lawyers of their obligations and confirm that they have verified the accuracy of any research done or case citations produced with the assistance of generative AI.

16. For litigants-in-person, many of them may not have the ability to verify the legal information provided by generative AI and may not be aware that such information is prone to error. If it appears that generative AI may have been used to prepare submissions or other litigation documents, JJOs should make enquiry with the litigant, and ask what checks for accuracy have been undertaken.

TWO GUIDING RULES

17. The above guiding principles can be succinctly summarised into the following two guiding rules –

Rule #1 : Do not delegate judicial functions to AI. Be
(on Input) mindful of what you enter/input into generative AI chatbots with respect to data security, confidentiality and privacy. Be aware of the risk

that anything entered/input will become information in the public domain; and

Rule #2 (on Output) : Be vigilant about the output generated by the AI chatbots, in particular factual accuracy, potential bias, infringement of intellectual property right, and use it at your own risk. You take responsibility for using AI and for the end product.

POTENTIAL USES

18. Generative AI may potentially be useful in the following tasks –

- (a) **Summarising information:** while AI tools are capable of summarising large bodies of text, care needs to be taken to ensure the summary is accurate and carries the same meaning as the original content;
- (b) **Speech/presentation writing:** AI tools can be used in planning a speech, producing an outline of potential speaking points and providing suggestions for topics to cover in a presentation;
- (c) **Legal translation;** and
- (d) **Administrative tasks:** while AI tools are useful in drafting e-mails/memoranda/letters, we need to be cautious that they can retain any data put to them (including names, e-mail addresses and so on), and could potentially disclose such information to a subsequent user.

19. As generative AI chatbots are limited by the date range, jurisdictional reach, and types of legal materials they can access, using them for research requires extra caution. Depending on the features of the AI model used and the underlying database, they may not be reliable as a complete substitute for conducting legal research using other means.

20. So long as the generative AI used generates text based on probability rather than an understanding of the nuances and context of any text, and does not have the ability to critically examine the patterns it identifies

in data, it can result in drawing inaccurate or biased conclusions and is therefore ill-suited to legal analysis. Unless there is a generative AI model with proven ability to protect confidential, restricted and private information and adequate built-in checking and verification mechanism to ensure accuracy and reliability, **using generative AI for legal analysis is not recommended.**

FURTHER UPDATES OF GUIDELINES

21. This is the first set of guidelines on the use of generative AI for judicial and administrative duties issued by the Judiciary, having referred to guidelines recently issued by courts in other jurisdictions as well as the Office of the Government Chief Information Officer and the Office of the Privacy Commissioner for Personal Data in Hong Kong.³ We will keep abreast of the latest developments of generative AI around the world and any new guidelines

³ These guidelines include (listed according to issue dates, starting from the most recent one):

- (a) AI: Model Personal Data Protection Framework published by the Office of the Privacy Commissioner for Personal Data on 11 June 2024 (https://www.pcpd.org.hk/english/resources_centre/publications/files/ai_protection_framework.pdf);
- (b) Interim principles and guidelines on the court's use of AI published by the Federal Court of Canada on 20 December 2023 (<https://www.fct-cf.gc.ca/en/pages/law-and-practice/artificial-intelligence>);
- (c) AI – Guidance for judicial office holders issued by the Courts and Tribunals Judiciary of England on 12 December 2023 (<https://www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf>);
- (d) Guidelines for use of generative AI in courts and tribunals issued by the Courts of New Zealand on 7 December 2023 (<https://www.courtsofnz.govt.nz/going-to-court/practice-directions/practice-guidelines/all-benches/guidelines-for-use-of-generative-artificial-intelligence-in-courts-and-tribunals/>);
- (e) Guidelines on ethical AI framework (Version 1.3) issued by the Office of the Government Chief Information Officer in August 2023 (https://www.ogcio.gov.hk/en/our_work/infrastructure/methodology/ethical_ai_framework/doc/Ethical_AI_Framework.pdf); and
- (f) The opinions on regulating and strengthening the applications of AI in the judicial fields published by the Supreme People's Court on 8 December 2022 (<https://www.court.gov.cn/fabu/xiangqing/382461.html>).

published by courts in other jurisdictions with a view to updating the guidelines as and when required.

ENQUIRIES

22. For enquiries on the above guidelines, please contact Mr Brian Chiu, Chief Judiciary Executive (Information Technology Office) at 2867 2669 or Mr Alex Lee, Senior Systems Manager (Information Technology Office (Technical)) at 2886 6895.

Judiciary Administration
July 2024

Some Common Terms Relating to Artificial Intelligence (“AI”)

Large Language Model (“LLM”):

LLMs are AI models which learn to predict the next best word or part of a word in a sentence having been trained on enormous quantities of text. Generative AI chatbots generally use LLMs to generate responses to “prompts”. Examples are ChatGPT and Bing Chat.

Generative Pre-trained Transformer (“GPT”):

A LLM based on the transformer architecture that generates text. It is first pre-trained to predict the next token in texts. After pre-training, GPT models can generate human-like text by repeatedly predicting the token that they would expect to follow. GPT models are usually fine-tuned to reduce hallucinations or harmful behaviour, or to format the output in a conversational format.

Token:

In natural language processing, a token is a unit of text that is processed by the AI, typically representing a word or a part of a word. However, a token does not have a fixed length in terms of characters or words. Instead, a token can vary based on the complexity of the language and the content.

Machine Learning:

A branch of AI that uses data and algorithms to imitate the way that humans learn, gradually improving accuracy. Through the use of statistical methods, algorithms are trained to make classifications or predictions, and to uncover key insights in data mining projects.

Deep Learning:

A function of AI that imitates the human brain by learning from how it structures and processes information to make decisions. Instead of relying on an algorithm that can only perform one specific task, this subset of machine learning can learn from unstructured data with supervision.

Data Mining:

The process of sorting through large data sets to identify patterns that can improve models or solve problems.

Natural Language Processing (“NLP”):

A type of AI that enables computers to understand spoken and written human language. NLP enables features like text and speech recognition on devices.

Technology Assisted Review (“TAR”):

AI tools used as part of the disclosure process to identify potentially relevant documents. In TAR a machine learning system is trained on data created by lawyers identifying relevant documents manually, then the tool uses the learned criteria to identify other similar documents from very large disclosure data sets.

Prompt:

Short instructions entered to a generative AI chatbot to obtain the desired answer/output.

Retrieval Augmented Generation (“RAG”):

RAG is a technique used in AI and natural language processing that aims to improve the quality of generated text by incorporating external information from various sources.

Hallucination:

An AI hallucination is when an AI system presents false information framed as the truth.

OpenAI:

OpenAI is an American AI company. It conducts AI research and has developed several AI models and services in the last decade, including GPT-3, ChatGPT and Dall-E.