



GUIDELINES FOR STAFF ON THE USE OF ONLINE AVAILABLE GENERATIVE ARTIFICIAL INTELLIGENCE TOOLS

1. PURPOSE OF THIS DOCUMENT

These guidelines are designed to help staff members⁽¹⁾ assess the risks and limitations of online available generative Artificial Intelligence (AI) tools and set conditions for their safe use in working activities of the Commission. They cover third party tools which are publicly available on the internet. The discussed risks and limitations are not necessarily relevant for internally developed generative AI tools from the Commission. Internal tools developed and/or controlled by the Commission will be assessed case by case under the existing corporate governance for IT systems.

As the technology of online available generative AI models evolves, so are the related risks, limitations, and the legal provisions. Therefore, this should be considered a living document that will be updated as needed, in particular taking into account the outcome of the negotiations of the Artificial Intelligence Act (proposal for a Regulation COM/2021/206).

2. WHAT IS GENERATIVE ARTIFICIAL INTELLIGENCE?

A new generation of online available Artificial Intelligence (AI) tools based on generative AI models have the potential of severely impacting our working methods. Some of the most well-known tools currently available online are ChatGPT, Dall-E, Midjourney, Bard AI, LaMDA, Aleph Alpha, Bloom, and Stable diffusion. Generative AI models are computer programs that are designed to create new content that resembles human-made output. Anyone could prompt such a model, typically using an online chatbot⁽²⁾, and ask it to write or summarise a text, generate a new image, suggest a graph to visualise a spreadsheet, etc.

The model works by learning patterns and characteristics from a large collection of data. When a user prompts the model, for example with a question or with an image, it generates a response that aligns as good as possible with its prior learning. This

(¹) These guidelines apply to all Commission staff, which means statutory staff (officials, temporary agents, and contract agents), seconded national experts and contractors or other service providers

(²) A chatbot is a computer program that uses artificial intelligence and natural language processing to simulate conversation with human users over the internet.

could give the impression that the model has human-like reasoning capacity. Generative AI models have great potential to increase the efficiency and improve the quality of our work. They could assist colleagues to write briefings, develop computer code for new software, or summarise citizen consultations, to mention just a few examples.

While it is important to explore the potential benefits of online available generative AI models for professional purposes, it is equally important to be aware of associated risks. Like any other information and communication tool used by staff in the Commission, online generative AI models should only be used when appropriate and safe.

3. RISKS AND LIMITATIONS OF GENERATIVE AI MODELS

First, there is a risk of unauthorised disclosure of information received in the line of duty. Any input provided to an online generative AI chatbot is transmitted to the provider of the chatbot, which can subsequently use this information to generate future outputs that could be disclosed to the public at large. By using generative AI tools, there is a risk that staff inadvertently make publicly available information for Commission use only, sensitive non-classified information, classified information, or personal data.

[Rule n°1] Staff must never share any information that is not already in the public domain, nor personal data, with an online available generative AI model ⁽³⁾.

Although the performance of generative AI models is constantly improving, staff must be aware that these models can produce biased and wrong answers to a user-prompt. This can be due to insufficient or biased training data, but also because of biases deliberately built in the algorithm. The creators of generative AI models are not always entirely transparent on the data and algorithms used. Hence, it is difficult to assess the reliability of an answer generated by such a model.

[Rule n°2] Staff should always critically assess any response produced by an online available generative AI model for potential biases and factually inaccurate information.

Lack of transparency on the origin of materials used for training generative AI models also raises concerns related to intellectual property rights, in particular copyright. It is possible that copyright protected information is used to train generative AI models, which may sometimes amount to infringing the copyright of the right-owners. This information could in turn end up being reproduced *verbatim* (or almost *verbatim*) in replies to a user-prompt, without any credit or reference to the source or the author. Moreover, it is currently not possible for generative AI models to properly list and credit the materials it is reproducing in its output, making it difficult to obtain the

⁽³⁾ We remind all staff that they are bound by Article 17 of the Staff Regulations, which forbids the unauthorized disclosure of any information received in the line of duty unless it is already public, even after leaving the service (see also Article 339 TFEU). Similarly, external contractors are bound by the security rules in their contract.

necessary authorisation from the right-owners, while model owners often waive any responsibility in this regard.

[Rule n°3] Staff should always critically assess whether the outputs of an online available generative AI model are not violating intellectual property rights, in particular copyright of third parties ⁽⁴⁾.

[Rule n°4] Staff shall never directly replicate the output of a generative AI model in public documents, such as the creation of Commission texts, notably legally binding ones.

Finally, online available generative AI models are known to show stability problems, such as long response times or unavailability of the service.

[Rule n°5] Staff should never rely on online available generative AI models for critical and time-sensitive processes.

4. CONCLUSION

In conclusion, while online available generative AI models can be explored to improve efficiency and quality of the Commission's work, staff should never disclose non-public information or personal data to the model. In addition, staff should always be vigilant when it comes to inherent risks and limitations. Any output of generative AI models must be checked for possible biases, inaccurate information, and copyright infringement.

⁽⁴⁾ For legal advice on intellectual property, Commission staff may always contact the Commission's Central Intellectual Property Service – EC-IPR@ec.europa.eu