Compass for the use of generative artificial intelligence (AI) in a work-related context

Version 2.0, 10th April 2024

Created by the AI Competence Network (headed by the Executive Group for Organisation and Security (MD-OS))

Stadt Wien

1. Compass for the use of generative artificial intelligence (AI) in a work-related context

Generative AI is a technology that uses „prompts", i.e. large amounts of data input, to generate content like texts, images, videos or audio contents. AI technologies offer enormous potential, but their practical application and dealing with their results also presents challenges.

In the spirit of digital humanism, we aim to make Vienna a capital of digital innovation that puts people first. We want to use generative AI conscientiously, and actively profit from the opportunities it offers. The present document is to help us use generative AI tools[1] responsibly in our work for the Vienna City Administration.

Our focus is on new tools that support the staff or teams and help them optimise work processes. The targeted implementation of AI technologies is part of our effort to promote innovation and complement human expertise so the staff can concentrate on more challenging tasks that require creative thinking and interpersonal skills.

For now, we have decided not to draw up a guideline with a catalogue of detailed requirements because AI is still developing dynamically. The coming months and years will see further innovation. What is more, legal regulations in this field are still pending and may require further adaptations once they are in place. This compass for the use of generative AI at work is subject to ongoing evaluation and will be updated if necessary.

1.1. Using generative AI for work

Generative AI available online may be used for work as this seems integral to a modern and efficient city administration. However, it is essential to observe existing framework conditions when using AI.[2]

It is important for us as City Administration staff to act responsibly and use our experience, common sense and professional expertise when using AI technologies. In this context, communication and

---

[1] Such as the language-based AI system "ChatGPT" or the AI image generator "Stable Diffusion", etc.

[2] Users should keep in mind decrees on electronic communication, data protection and IT security, such as MD-OS-338293/2019 „Internet und elektronische Kommunikation; offizielle Dienststellen-Postfächer" (Internet and electronic communication, official inboxes of departments), MDK – 2008102-2022-1 „Datenschutz im Magistrat der Stadt Wien" (Data protection in the Vienna City Administration) or MD-OS 1416182-2023 "Sicherheit in der Informations- und Kommunikationstechnologie" (Security in information and communication technology), as well as the City of Vienna's Corporate Identity Manual "Das Erscheinungsbild der Stadt Wien" and the guidelines on gender-neutral communication in "Leitfaden für geschlechtergerechtes Formulieren und diskriminierungsfreie Bildsprache" (Guidelines for non-discriminatory language and images).

coordination are key, both within our team and with the executive level, so we can jointly identify the most suitable areas of application. After all, AI holds more potential benefits for some areas of work than for others.

<u>As always in working life, we must observe the following principles as we proceed:</u>

a. Legality (e.g. data protection, copyright, etc.)
b. Economy and efficiency (cost/benefit ratio) and expediency (promoting innovation, improving performance)
c. Compliance with other legal and ethical standards

This will allow us to use generative AI as an effective support tool in our field of work.

1.2. What does this basic principle imply when using AI tools for work?

1.2.1. Human users remain responsible when using content generated by AI.

As we use generative AI tools, it is essential to ensure the results of AI undergo critical review and are checked for content accuracy and ethical acceptability (e.g. equal treatment and non-discrimination). This is vital because text-based AI tools are primarily language models (rather than "knowledge libraries"), and have been known to make up content (- "hallucinate"). The responsibility for content we choose to reuse always rests with us, the (human) users.

1.2.2. Generative AI can help support, simplify and speed up work processes in public administration.

The quality of the results obtained with generative AI tools depends on the quality of the prompts, i.e. input, provided by the users. It often improves with further question-answer loops between the user and the AI system. When using generative AI tools, we should also make sure we read and follow the instructions in the user manual for the tool.

1.2.3. It is important to observe legal regulations when dealing with personal data or information subject to official secrecy.

Users who enter content into a generative AI tool should follow the same principles that apply to the internet or social media. This means they must observe all relevant legal regulations when dealing with personal data (including images), classified internal information, business or company secrets, or information subject to official secrecy.

1.2.4. If content has been generated with the help of AI, this fact must be pointed out clearly:

"The content of this text/image was generated with the help of an artificial intelligence tool and has undergone human review."

If a generative AI tool has been used to generate text, images etc., this fact must at least be disclosed to the outside, e.g. in a footnote to the respective text, in the image caption or the e-mail signature. However, adding a note disclosing the use of AI does not mean the person who created or signed the document is no longer responsible for the content. There is no need to disclose AI use, however, if you have only modified existing content rather than adding new content, as is the case with spell checkers, translation programmes, text summaries, e-mail text elements, or photos cropped with the help of AI.

1.3. What does this basic principle imply if my work involves dealing with results that (may) haven been generated with AI tools?

As with any information from external sources (e.g. social media, internet, etc.), AI-generated content must undergo critical review, and you have to verify its accuracy before using the content as a basis for decision-making or work processes.

In particular, AI-generated "deepfakes" – images, audio or video recordings that seem real but have been manipulated with AI – and their rapid dissemination on social media present considerable challenges. They can lead to disinformation and its further consequences, potentially compromising the credibility of the City Administration.[3]

For this reason, it is vital to check the underlying facts so you can tell the difference between real and fake content, especially fake images that may have been generated with AI tools.

A number of further criteria[4] can help you identify indicators of authenticity, e.g.:

a. Does the context of the image look conclusive?
b. Do all components of the image make sense in this context?
c. Are there any obvious mistakes, or do elements of the image look as if they "don't fit properly", e.g. body parts?
d. Do you notice any blurred or merged transitions (in particular between a person and the image background)?
e. Do you notice any inconclusive reactions of people shown in the image, or any other unusual appearances?

---

[3] *https://www.wien.gv.at/medien/fake-news/*

[4] cf. *https://apa.at/faktencheck/ki-generierte-bilder-haeufig-geteilt/*

1.4. KITT – Hints and hacks for working with AI in practice

The Vienna City Administration has launched a platform called KITT where staff can exchange information and learn more about artificial intelligence. KITT requires no special (technical) know-how and is open to everyone working for the City Administration.

The recurrent KITT online focus programme has three pillars:

a) Introduction, inspiration and challenge (held once a month, for newcomers)

b) Exchange of experience and update on current developments (once a month, helps you keep an overview)

c) Prompt Together (cooperative community event, held twice a month to practise effective prompting)

To receive further info and announcements on KITT, just send an e-mail with the caption "Terminverteiler" to kitt@ma53.wien.gv.at. You can also join the KITT groups on the City of Vienna Intranet and in Signal Messenger.

Title	Compass for the use of generative artificial intelligence (AI)
Version	2.0
As of	10th April 2024
As amended	Version 1.0
Published by	Chief Executive Office of the City of Vienna Executive Group for Organisation and Security Organisation Group (MD-OS/ORG)
Contributions by	AI Competence Network
Document classification	Freely available
Enclosures	-