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

Version 3.0, 1 July 2025

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



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

Generative Al is a technology that uses "prompts", i.e. large amounts of data input, to generate content like texts, images, videos, or audio contents. Al 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<sup>1</sup> responsibly in our work for the Vienna City Administration.

Our focus is on new tools that support staff or teams and help them optimise work processes. The targeted implementation of Al 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, the legal framework for the use of AI is set out in the AI Act and is constantly specified and improved. This compass for the use of generative AI at work is subject to ongoing evaluation and will be updated if necessary.

### 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<sup>2</sup>.

Reference is also made to the City of Vienna's Corporate Identity Manual "<u>Das Erscheinungsbild der Stadt Wien"</u> and the guidelines on gender-neutral communication and non-discriminatory language and images "<u>Leitfaden für geschlechtergerechtes Formulieren und diskriminierungsfreie Bildsprache"</u>.

Al Compass 1 Juli 2025 Page 2 of 6

<sup>&</sup>lt;sup>1</sup> Such as the language-based AI system "ChatGPT" or the AI image generator "Stable Diffusion", etc.

<sup>&</sup>lt;sup>2</sup> Users should keep in mind decrees such as MD-OS-338293/2019 "Internet und elektronische Kommunikation; offizielle Dienststellen-Postfächer" (Internet and electronic communication, official inboxes of departments) and MDK – 2008102-2022-1 "Datenschutz im Magistrat der Stadt Wien" (Data protection in the Vienna City Administration), and MD-OS 1416182-2023 "Sicherheit in der Informations- und Kommunikationstechnologie" (Security in information and communication technology).

### **Basic principle**

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 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.

As always in working life, we must observe the following principles:

- 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<sup>3</sup>.

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

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

### 2.1. Human users remain responsible when using content generated by Al

As we use generative AI tools, it is essential to ensure that the results of AI undergo critical review and are checked for content accuracy and ethical acceptability (e.g. equal treatment and non-discrimination). The responsibility for content we choose to reuse always rests with us, the human users.

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

Generative AI tools have the potential to transform the services provided by the City of Vienna, make them more efficient, and customise them to the personal needs of citizens. Their benefits also lie in their ability to analyse large amounts of data and recognise patterns, which makes generative AI useful in many areas. Without appropriate control, generative AI also harbours risks for individuals and society. Therefore, we need to choose a controlled, responsible and ethical approach when dealing with generative AI.

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 manuals.

Al Compass 1 Juli 2025 Page 3 of 6

<sup>&</sup>lt;sup>3</sup> Ethical standards mean, in particular, fairness (i.e. that Al-supported decisions should not discriminate or disadvantage) as well as transparency and explainability (i.e. that the use of Al should be made clear and comprehensible). This is based on the principle of digital humanism, which is a key pillar of the City of Vienna's Al Strategy.

### 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 that 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.

### Different approaches depending on the AI tool used

- a. If you use AI tools that are freely available on the internet, you are only allowed to enter freely available (public) data there.
- b. Based on a data protection impact assessment carried out by MA 01 on the operating environment of the "Vienna AI platform" (formerly "My ChatGPT platform"), it is, for the time being, permitted to enter data into this AI tool that is classified as "restricted" or "confidential" according to security classification levels<sup>4</sup>. However, you must be aware that the data protection impact assessment only covers general data protection issues and risks, and you must still adhere to data protection regulations when entering personal data into this AI tool based on your specific processing activities. The impact assessment does not cover the legality of data input that is subject to further confidentiality requirements pursuant to legal regulations that only affect individual departments in the performance of their tasks and must be assessed by the departments themselves.
- c. If you use any other AI applications approved for use in the Vienna City Administration, you must comply with the appropriate requirements set out as part of the relevant approval process in accordance with the AI Act.

### 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 email 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, email text elements, or photos cropped with the help of AI.

Al Compass 1 Juli 2025 Page 4 of 6

<sup>&</sup>lt;sup>4</sup> Supplement to the decree "Sicherheit in der Informations- und Kommunikationstechnologie" (Security in information and communication technology) of 22 November 2023 (MD-OS-1416182-2023).

# 3. What does this basic principle imply if my work involves dealing with results that (may) have been generated with AI?

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

### 3.1. Hallucinations

It is important to note that while generative AI is powerful, it is not always perfect. Therefore, human review helps to ensure quality and accuracy.

This is vital because text-based AI tools are primarily language models and have been known to make up content (i.e. "hallucinate").

### 3.2. Fakes

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.<sup>5</sup>

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 can help you identify indicators of authenticity, e.q.:

- 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 do not 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 unusual appearances?

### 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) knowledge and is open to everyone working for the City Administration.

Al Compass 1 Juli 2025 Page 5 of 6

<sup>&</sup>lt;sup>5</sup> https://www.wien.gv.at/medien/fake-news/

<sup>&</sup>lt;sup>6</sup> cf. https://apa.at/faktencheck/ki-generierte-bilder-haeufig-geteilt/.

### KITT series of online events:

- a. Introduction, inspiration and challenge (held once a month, for newcomers)
- b. Exchange of experiences and update on current developments (once a month, helps you keep an overview)
- c. Prompt together (twice a month to practise effective prompting)

Send an email with the subject line "Terminverteiler" to <a href="kitt@ma53.wien.gv.at">kitt@ma53.wien.gv.at</a> to be added to the mailing list. You can also join the KITT groups on the City of Vienna Intranet and in Signal Messenger.

### 5. Mandatory e-learning courses

The Personnel Training Academy of the City of Vienna provides the following e-learning courses. It is mandatory for City Administration staff to complete these courses once:

- a. Al artificial intelligence
- b. Fake news

Title	Al Compass
	<del>                                     </del>
Version	3.0
As of	1 July 2025
As amended	Version 2.0
Published by	Vienna City Administration
	Executive Group for Organisation and Security
	Organisation Group
	(MD-OS/ORG)
Contributions by	Al Competence Network
Document	Freely available
classification	
Enclosures	-

Al Compass 1 Juli 2025 Page 6 of 6