Guidelines for Responsible Use of Generative Artificial Intelligence in the Warsaw City Hall

The guide was developed by City of Warsaw in cooperation with the SWPS University, EGO (Evaluation for Government Organizations) S.C., and the Center for Ethics of Technology at the Humanites Institute

Warsaw

patronage of

Ministry of Digital Affairs
Republic of Poland

Guidelines for Responsible Use of Generative Artificial Intelligence in the Warsaw City Hall

dictionary → 3

introduction → 5

guiding principles → 7

how to use generative artificial intelligence responsibly → 10

annex — recommendations for those procuring or implementing GenAI Tools at the Warsaw City Hall → 23

application examples → 25

DICTIONARY

Artificial intelligence

Artificial intelligence is a general term for various programming techniques that allow machines to perform tasks that normally require the intelligence of humans or other living beings. These systems can operate more or less autonomously and can, for example, recognize patterns, make predictions, give recommendations or make decisions. These technologies rely on data processing, model building and algorithm development to automate processes such as data collection, processing and analysis.

Generative artificial intelligence

Generative artificial intelligence is a type of artificial intelligence that learns from available data and then creates new content, such as texts, images or codes. Unlike traditional AI systems that only analyze data, GenAI can generate new content. It is characterized by “creativity” and adaptability to a variety of tasks. It can create, for example, personalized responses to questions from officials or residents, assist in the preparation of reports, or in the design of urban spaces.

GenAI tools

GenAI tools is a software that uses generative artificial intelligence to perform a specific task. This category includes a wide range of tools, from the freely available ChatGPT, which can perform almost any task, to a specialized in-house system that was created for a specific group of tasks, such as supporting the work of the legal office at the City Hall.

Prompt

Prompt is a command or instruction used to generate a response or content by an artificial intelligence model, such as in generative AI systems.

Journal (GenAI use)

Journal (GenAI use) is a place designed to record the steps involved in applying generative artificial intelligence technology at work.

INTRODUCTION

As a modern and dynamic city, Warsaw is committed to innovation that can realistically improve the quality of life for its residents. One of the tools to achieve this goal is generative artificial intelligence (GenAI).

The technology offers tremendous opportunities to create new solutions and improve public services. Warsaw, acting as inspiration for other cities, wants to use this potential in a responsible and thoughtful way.

Generative artificial intelligence is a groundbreaking technology that can significantly contribute to the “We generate innovations” program whose goals are set forth in the #Warszawa2030 Strategy.

The Warsaw City Hall recognizes the enormous potential of this technology, but also the challenges it comes with. Therefore, we are taking steps to create rules for the responsible use of GenAI. This document provides a framework for using this technology in public administration, ensuring the protection of residents' rights while building trust in public institutions.

Our recommendations include:

- supporting City Hall employees by utilizing GenAI tools,
- acting for the common good,
- responsible use of technology,
- risk management,
- data protection,
- transparency of operations and understandable decisions,
- continuous improvement of competencies.

We want GenAI to support the improvement of public services and the daily lives of people living in Warsaw, at the same time ensuring that our actions are in accordance with ethical and legal principles. The document is a signpost on how to use GenAI responsibly for every employee at the Warsaw City Hall.

This guide will be updated on an ongoing basis to better meet the needs of those working in the City Hall and to take into account the opinions of Warsaw residents. The changes will also be aligned with new legislation and technological advances.

Warsaw aspires to be the leader in the responsible use of modern technology. We want to inspire and contribute to the creation of a smart, sustainable and inclusive city of the future, without forgetting the role of residents' art and creativity.

The document demonstrates our commitment to responsible innovation and adaptation to the rapidly changing technological reality — always keeping in mind the well-being of Varsovians.

guiding principles

The recommendations contained herein for the responsible use of generative artificial intelligence by City Hall employees take into account EU guidelines[1] in terms of: respect for human autonomy, prevention of harm, fairness, and ability to explain results.

Using these guidelines as a corner stone, rules have been designed that will facilitate proper use of generative AI at work:

1. human agency and oversight — refers to the principle of respect for human autonomy;

2. technical robustness and safety — refers to the principle of preventing harm;

3. privacy and data governance — refers to the principle of preventing harm;

4. transparency — refers to the principle of explainability;

5. diversity, non-discrimination and fairness — refers to the principle of justice;

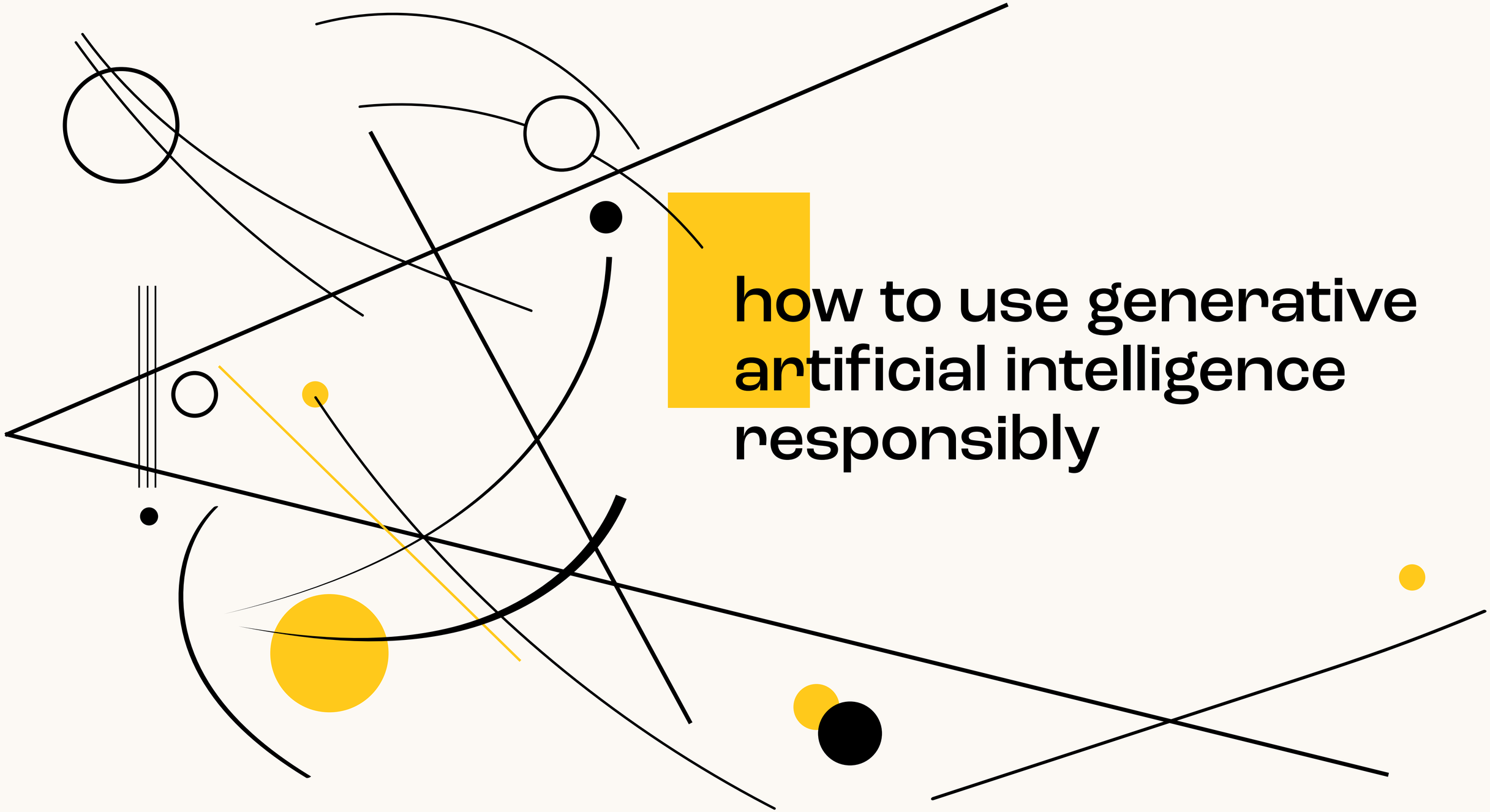
6. societal and environmental well-being — refers to the principles of justice and preventing harm;

7. accountability — refers to the principle of justice, and is complementary to all the previous requirements.

The aforementioned principles are also cited in Recital 27 of the Preamble to the Act of the European Parliament and of the Council on the establishment of harmonized provisions on artificial intelligence (AI Act)[2]. This act, together with the Ethics Guidelines, are the basic normative acts describing the concept of trustworthy artificial intelligence (trustworthy AI) accepted in the European Union.

The standards written in this document also take into account the principles of the European Declaration on Digital Rights and Principles for the Digital Decade[3], namely:

- → putting people at the centre of the digital transformation,
- → solidarity and inclusion,
- → freedom of choice,
- → participation in the digital public space,
- → safety, security and empowerment,
- → sustainability.



how to use generative artificial intelligence responsibly

How to use generative artificial intelligence responsibly

When using generative AI tools in City Hall,
one should keep in mind a few rules for its application:

1. GenAI is about support, not replacement → 12
2. We use GenAI for the common good → 14
3. We take responsibility for the way we use GenAI → 15
4. We can evaluate the risks and we follow the rules for the safe use of GenAI → 17
5. We care about data security → 19
6. We use GenAI tools in an open and transparent manner → 20
7. We gain knowledge about the responsible use of GenAI → 21

1. GenAI is about support, not replacement

Why it is important: The task of generative artificial intelligence is to support our work so that it is more efficient. We decide when to use these tools. We also have control over what is created using GenAI. We can change or correct the content created by these tools at any time.

Key indications:

→ GenAI can be a useful tool, but it is up to the individual to decide about its application. The tools are meant to streamline our work and support us in helping residents, city users, and each other.

→ Technology can support our skills, but it should not lead to mechanical completion of tasks — especially those requiring attention, specialized competence or creativity.

→ We control the manner in which GenAI tools are used — if they do not meet our needs, we can stop the AI-supported work at any time.

→ Innovation is not the end goal in itself. AI is here to help us, but it is not a substitute for contact with another human being — this is true for both the public, as well as those working in the City Hall.

→ We use GenAI tools when, according to our judgment, they are needed. For example, if it takes us as much time to verify a recommendation generated by an artificial intelligence tool as it does to make a decision on our own, we consider whether AI adds any real value to the process.

Please remember:

- → AI is only our assistant — we make the final decision.
- → We use GenAI when it actually supports our work.
- → When in doubt, we check the results of generative artificial intelligence. We trust our competence and knowledge above all.

see application example

2. We use GenAI for the common good

Why it is important: Our goal is to support Warsaw residents and users of the city and their environment. The application of GenAI in one area may have implications in another area as well. Therefore, we are trying to consider the impact of generative artificial intelligence in various aspects of city operations.

Key indications:

- The City Hall is a public institution, so as its employees, we have a special obligation to use GenAI support responsibly, with the well-being of the city's users in mind.
- When using GenAI tools, we make sure that their impact is positive socially and environmentally. We take into account not only their impact on the quality of life of individuals, but also on communities and the environment.
- We make sure that the content generated by GenAI does not discriminate against any social groups.

Please remember:

- The needs of the city's residents and users come first — not technology.
- It's not just the “here and now” that matters — the long-term impact is also important.
- We also check the generated content to avoid discrimination.

see application example

3. We take responsibility for the way we use GenAI

Why it is important: Responsibility for the use of GenAI applies to the entire organization, not only one person. We belong to a large public institution and following procedures helps us to use technology responsibly. We check all AI recommendations, and when in doubt, we turn to experts.

Key indications:

→ Each person using GenAI is individually responsible for taken actions — regardless of position or function.

→ Using GenAI in our work, we use only systems approved by the City Hall — and their list is kept up to date.

→ We have to remember that, as with any other tool, we are bound by the law (e.g., RODO) and internal regulations (e.g., presidential ordinances).

→ The City Hall has a department specialized in issues of responsible use of generative artificial intelligence. Employees can easily access it. It provides assistance in the safe, legal, and ethical use of this technology.

→ When a new GenAI tool is introduced at the City Hall, procedures are created that indicate its proper use. Users' common sense and limited trust in artificial intelligence, on the other hand, help to apply it responsibly.

→ If the answers generated by GenAI provide valuable input to our decision, they have to be reviewed and approved by us. This is especially important when they carry legal implications for residents and users of the city. We check them carefully before publication — the so-called "second look rule" applies, which involves a second person verifying the results.

→ In situations where an internal GenAI tool makes mistakes, malfunctions or acts abnormally, we document and report the event to the administrator of the system in question accordingly.

→ Failure of GenAI tools does not paralyze our work — we have a plan of action in case the system stops functioning or generates errors.

→ The use of generative artificial intelligence tools in high-risk areas is recorded.

Please remember:

- → We only use GenAI according to procedures and do not accept its prompts without consideration.
- → If we have doubts, we ask experts or look for reliable sources (such as scientific articles).
- → We report noticed errors in order to verify them and pass on the error information to other users.

see application example

4. We can evaluate the risks and we follow the rules for the safe use of GenAI

Why it is important: The dangers associated with GenAI can vary and are not always obvious. Cybersecurity compliance alone does not ensure that the results generated by AI tools will comply with applicable laws or accepted ethical principles. It is also worth bearing in mind the context of the use of artificial intelligence.

Key indications:

→ Risks associated with the use of GenAI can occur in a number of areas, including information leakage, the image of the institution, possible harm to residents and users of the city, or impact on the well-being of citizens and the environment.

→ In a situation where a new GenAI tool has been developed for the City Hall, ethical and legal requirements are included in the contract specifications and are treated equally with system security requirements.

→ If we have any doubts about the possible risks of using GenAI in a particular case, we consult with our superiors and the department responsible for using AI in the City Hall.

→ In our work, we use GenAI tools approved by the City Hall. Their list will be updated on a regular basis.

→ We enter business documents using the business account only to those GenAI tools that the City Hall has explicitly classified as secure and suitable for this type of use.

Please remember:

→ We know about the various risks associated with the use of GenAI.

→ We treat these standards and the applicable laws on a par with safety, efficiency or productivity requirements.

see application example

5. We care about data security

Why it is important: We approach data with care, so that it is properly protected. When working with GenAI tools, as with other tools, we use only the data we have the right to use. We are also committed to ensuring that the city's residents and users know how to keep their data safe when dealing with GenAI tools provided by the City Hall.

Key indications:

- → When using GenAI, we keep in mind the principles of information security.
- → We specifically secure protected data — relevant instructions specify which data must never be entered into GenAI tools.
- → When working with data authorized for processing in GenAI tools, we ensure compliance with procedures and exercise caution.
- → In accordance with RODO, we are granted access to and the ability to process data to the extent necessary to perform our duties, and the actions we take are recorded in our activity log.

Please remember:

→ We secure not only the data entered into the system, but also the data generated by it.

→ The system administrator can check what data was entered into the internal GenAI tool offered by the City Hall.

see application example

6. We use GenAI tools in an open and transparent manner

Why it is important: The use of generative artificial intelligence can raise doubts and concerns, so we report on its use at work directly. It is worth remembering that by hiding the use of GenAI, we risk increasing residents' distrust of the technology and the City Hall's actions.

Key indications:

→ After getting a response from GenAI, we ask the tool about the process of arriving at the result, i.e. we ask about the source of the data from which the system generated the content.

→ We take care of procedure transparency — if necessary, we inform stakeholders (residents, users of the city and employees of the City Hall) about the use of GenAI in a particular case.

→ We have a journal where we record the use GenAI for support in decision-making. This ensures that the GenAI tools used at the City Hall can be reliably audited.

→ When GenAI tools are used to create audiovisual content, we always inform about their use by labeling it with the caption "The image was generated with the assistance of GenAI."

see application example

Please remember:

→ In the prompt, we put in a command for describing the way the generated content has been achieved.

→ When creating audiovisual content, we report on the use of AI.

→ We register the use of GenAI tools in decision-making by recording them in the journal.

7. We gain knowledge about the responsible use of GenAI

Why it is important: GenAI technology is developing rapidly. With new opportunities come previously unknown risks. Once acquired knowledge is not enough to safely use GenAI tools. It should be regularly supplemented with new knowledge and competencies in this area.

Key indications:

→ We participate in training courses offered by the City Hall, ensuring that we acquire the competencies necessary for the safe and responsible use of GenAI at the City Hall.

→ We use guides and webinars on responsible use of GenAI in order to level standards of knowledge among employees of different specialties.

→ We can participate in periodic meetings, during which City Hall staff and outside experts present good practices on the use of AI and current issues on AI security, ethics and law.

→ We know and understand the risks involved in using GenAI tools in our work, and we regularly supplement our knowledge of how to use these tools safely.

see application example

Please remember:

→ The City Hall provides us with accredited learning materials on GenAI tools.

→ All employees have access to knowledge regarding this technology — not only IT staff.

→ We regularly acquire new skills to use GenAI safely and responsibly.

The document outlining recommendations for the responsible use of generative artificial intelligence in the Warsaw City Hall should be regarded as active and dynamic. It will change in response to the future needs of those who use it, as well as to technological advances.

Therefore, we encourage you to regularly review its content and make any comments or suggestions that may contribute to its improvement.

Please read the WARSAW GenAI CODE, which is a compilation of the most important information presented in this publication.

annex
—
recommendations for those procuring or implementing GenAI tools at the Warsaw City Hall

- When considering the implementation of a particular GenAI system at the City Hall, we take into account not only the immediate impact of its application, but also the broader context (e.g., an external provider of AI tools should meet the requirements of the law and the ethics of artificial intelligence, including respect for basic rights).

- We provide easy access to GenAI's systems for all authorized individuals.

- When designing tools, we consult solutions with representatives of various stakeholders and take their opinions into account.

- When ordering a new GenAI tool, one has to require guides and instructions from the vendor on how to use the solution responsibly.

- Before the tools of GenAI systems are put into use in the City Hall, tests must be carried out for technical integrity and proper functioning, as well as for compliance with applicable regulations, internal ordinances.

- The data resources we use while working with the various GenAI tools are regularly checked for integrity, security, and updates by authorized specialists.

- When ordering tools, we evaluate the level on which the data entered into the system may become sensitive, e.g. due to correlation.

- In our internal tools, we ensure that the data used in GenAI systems comes from reliable sources.

- When performing a specific task, we check if the GenAI retrieves data from the City Hall's database and if the data is up-to-date.

- Materials created with GenAI's support are regularly checked to quickly detect any anomalies (such as issues of overrepresentation or underrepresentation of specific social groups).

- The functioning of GenAI tools is verified by people with expertise in data analysis and experts from the City Hall who can explain a specific action, if necessary.

- Answers generated by GenAI must not lead to inconsistencies with the law and City Hall's internal regulations — it is our job to check and modify them accordingly.

application examples

GenAI is about support, not replacement

Example: Moderation of social media content of the Warsaw City Hall.

Context

Warsaw City Hall manages official social media profiles on which residents actively comment and ask questions. Communication team has to moderate hundreds of comments daily — a very time-consuming activity that might lead to overlooking important content.

The way GenAI supports the process

1. A language model is able to analyze and categorize comments as, for example: potential abuse / neutral comments / valuable remarks.
2. A language model is able to create a priority list of comments that require attention.
3. GenAI is able to suggest templates for answers for frequently asked questions (FAQs).

Sample prompt:

You are acting as a social media moderation assistant for Warsaw City Hall. Your task is to analyze and categorize user comments. For each comment:

1. Select a category:
   a. Potential abuse (for example, hate speech or spam),
   b. Neutral comment,
   c. Valuable remark (constructive criticism, improvement suggestion).
2. Set the priority level (high/medium/low) based on the urgency and potential influence.
3. If the question is frequently asked, suggest an answer template.
4. Identify key words or topics that might be valuable for future analysis.

Analyze the comment below and present the results in a structured format:

[PASTE THE COMMENT THAT REQUIRES ANALYSIS HERE]

Please make sure that:

→ You take full responsibility for moderating the content. You are the one making the decision based on priority lists, you teach and improve the system by correcting erroneous categorizations, you identify topics for public consultation based on valuable remarks, and you conduct empathetic communication by personally formulating responses on sensitive issues.

→ You check that the content analyzed by GenAI does not contain sensitive data, such as users' personal data.

→ You document the analysis process in case the decision you made needs to be clarified.

back to guidelines

We use GenAI for the common good

Example 1: Adjusting the message about the new garbage collection rules for different audiences.

Context

Your task is to inform residents about the new garbage collection rules. You realize that different groups of residents may need different forms of communication, but you lack experience in creating differentiated messages.

The way GenAI supports the process

1. Prepare basic information about the new garbage collection rules.
2. Identify the groups for which you want to tailor the message (for example: seniors, families with young children, foreigners, etc.).
3. Use generative AI to create different versions of the message to different audiences.

Sample prompt:

Create a message about the new garbage collection rules in Warsaw for [TARGET GROUP]. The new rules are:

1. Segregation into 5 fractions: paper, glass, metals and plastics, biowaste, mixed waste
2. Collection of segregated waste every two weeks
3. Collection of mixed waste every week
4. The necessity of washing biowaste containers

The message should be:

- Simple and understandable
- Taking into account the specific needs of the target group
- No longer than 150 words
- Containing information on where to find more details

[TARGET GROUP] is: [seniors / families with young children / foreigners who are not fluent in Polish]

Please make sure that:

- → All generated messages contain correct information.
- → All versions of the message convey the same key information — despite differences in form.
- → Evaluate whether the messages actually meet the needs of each group.
- → Remember that as a human being you take full responsibility for the content — choose the best versions and modify them, if necessary, to better suit your expectations.

back to guidelines

Example 2: As a result of public consultations on the construction of a new road in the district, you receive 500 comments from residents. You use the GenAI system to organize and synthesize the collected postulates.

Context

Your task is to prepare a summary of the public consultations on the route of a new road in the district. It is crucial that all the insights of the residents are taken into account, and that the synthesis of the material does not leave out any important information. The use of GenAI can significantly improve the process, while requiring special attention to fairness and non-discrimination.

The way GenAI supports the process

1. Generative AI can quickly process hundreds of comments from residents, identifying main themes and concerns.
2. The GenAI system can group similar opinions, facilitating their analysis and presentation.
3. Generative artificial intelligence can evaluate residents' overall attitude toward various aspects of the project.
4. Based on the analysis, GenAI can create a draft summary, which you will verify and refine.

Please make sure that:

- → The system understands a variety of speech styles, dialects and perspectives.
- → You randomly check the original comments and compare them with GenAI's summary.
- → The system does not overemphasize the most popular opinions at the expense of less numerous, but important ones.
- → Clearly communicate to your audience that GenAI tools were used in the process, thus maintaining transparency.

Please remember:

GenAI supports the analysis, but you are responsible for the integrity of the decision. Make sure that every voice is taken into account and the report does not discriminate against anyone. Critically evaluate AI results using your knowledge of the community. This is the only way to ensure that the final summary is fair and representative of all residents.

back to guidelines

We take responsibility for the way we use GenAI

Example: Responsibility for using GenAI in creating official reports.

Context

Your task is to prepare weekly reports based on data on the use of city resources (for example: equipment, vehicles, office supplies). These reports are key to resource allocation decisions for the following week.

Just before the weekly meeting, you discover the capabilities of the generative AI tools offered by the City Hall and start using the language model to analyze data and create reports, which has greatly accelerated your work (the data is not confidential or sensitive). Before the meeting with management, it turns out that the latest report generated by GenAI contains errors. Some of the numbers do not agree with the source data, and the conclusions are illogical in places.

You are facing a difficult choice:

a. you can use the report, hoping that no one will notice the errors,
b. you inform your superiors of the error, explaining that it was AI that made the mistake,
c. you take full responsibility for the mistake and take corrective action.

Please make sure that:

→ In light of the rule described, the correct approach would be to inform your superiors of the inaccuracies discovered in the report (you choose option c).

→ You have a record of the procedure, which will help you understand at what point the error occurred and how to avoid a similar situation in the future.

→ As a human being, you take full responsibility for the results of GenAI's work, which in practice means that you check them at least randomly.

back to guidelines

We can evaluate the risks and we follow the rules for the safe use of GenAI

Example: Detecting and reporting an error in GenAI work.

Context

Your team uses the GenAI model to initially categorize and respond to residents' inquiries sent electronically. You notice the model making a mistake in interpreting questions about the new waste segregation program — the system consistently gives outdated information about the bulky waste collection schedule, even though the program was recently updated.

Please make sure that:

- → You conduct several trials, asking the model different variants of bulky waste collection questions to confirm the systematic nature of the error.
- → You document each attempt by recording the exact question and GenAI's answer.
- → You prepare a short report describing the problem you noticed and the results of the tests and send it to the department responsible for using GenAI tools at work.

back to guidelines

We care about data security

Example: Analyzing data with GenAI tools.

Context

You work in the Department of Economic Development of the Warsaw City Hall. Your task is to analyze the data of companies registered in Warsaw to identify potential beneficiaries of the new support program for small and medium-sized enterprises (SMEs). You decide to use the GenAI tool to analyze publicly available data from the National Court Register (KRS).

Remember to use publicly available KRS data that does not contain sensitive information. If the City Hall has isolated models (which do not require connection to external servers), it will be possible to securely process internal and sensitive data as well.

Sample prompt:

Your task is to create a list of companies that potentially qualify for the SME support program. Using data from the KRS, analyze information on company size, industry and registration date to identify relevant entities.

Eligibility criteria are:

- Company registered in Warsaw
- Employment does not exceed 250 employees
- The company operates in the service or manufacturing sector (excluding trade)
- The company has been registered for at least 2 years

Generate a list of companies that meet the above criteria, taking into account the following data:

- Name of the company
- KRS number
- Registered office address
- Date of registration
- Main activity (PKD code)

Please make sure that:

- The data used by the tool is correct and up-to-date.
- You understand how GenAI arrived at such results, that is, by what rules it pulled the list of people from the general pool. Typically, the model will create a script that it will use to analyze the database. Save it, even if you are not familiar with programming.
- You randomly check the results — they should agree with your calculations (you don't have to check all the results — if it had to be done, it would be pointless to use AI — but try to always randomly check the results). Remember, as a human being, you take full responsibility for the result of GenAI's work.

back to guidelines

We use GenAI tools in an open and transparent manner

Example: Creating images using generative AI.

Context

You work in the Waste Management Department and your task is to prepare informational materials about the new rules of garbage segregation in Warsaw. You decide to use the generative AI tool to create images illustrating these rules.

To generate images, you can use tools which will create illustrations that comply with the new segregation rules. However, remember to be transparent.

Sample prompt:

Create an image to illustrate the new rules for garbage segregation and collection.

The image should illustrate the segregation of waste into 5 fractions:

1. Blue container — for paper
2. Green container — for glass
3. Yellow container — for metals and plastics
4. Brown container — for biowaste
5. Black container — for mixed waste

Style: simple and clear information poster.

Result generated by MS Copilot program:

The image was generated with the assistance of GenAI

If you find the result unsatisfactory (the image shows only 4 containers, and one of them is orange, not brown), make another attempt by changing the prompt. When you get the result ready to publish, remember to sign it: “The image was generated with the assistance of GenAI”.

Please make sure that:

→ The image generated by GenAI is of adequate quality, and the content is not questionable.

→ Information about the use of GenAI tools is visible.

back to guidelines

We gain knowledge about the responsible use of GenAI

Example: Testing a new GenAI tool that has greater capabilities than official systems.

Context

As someone interested in artificial intelligence, you come across an interesting article about a new AI tool for document formatting. There is no such tool in the catalog of tools accepted and installed by the City Hall, but you want to test it, because it can help you save a lot of time.

Please make sure that:

→ You read the terms of use and privacy policy (whether the tool is suitable for professional use).

→ You remember to use AI tools at work only on business devices and accounts. Before you ask for appropriate approval, evaluate its practical usefulness.

→ You do not use protected data or confidential documents for testing. For example, you use documents and data that are publicly available on the Internet for testing.

→ You document the tests, describing your actions and results.

Please remember:

These actions will help create a report that you can forward to the department in charge of responsible use of AI to decide whether to allow the use of this tool in the work of the City Hall. This will ensure that the tool catalog is not strictly defined once and for all.

back to guidelines