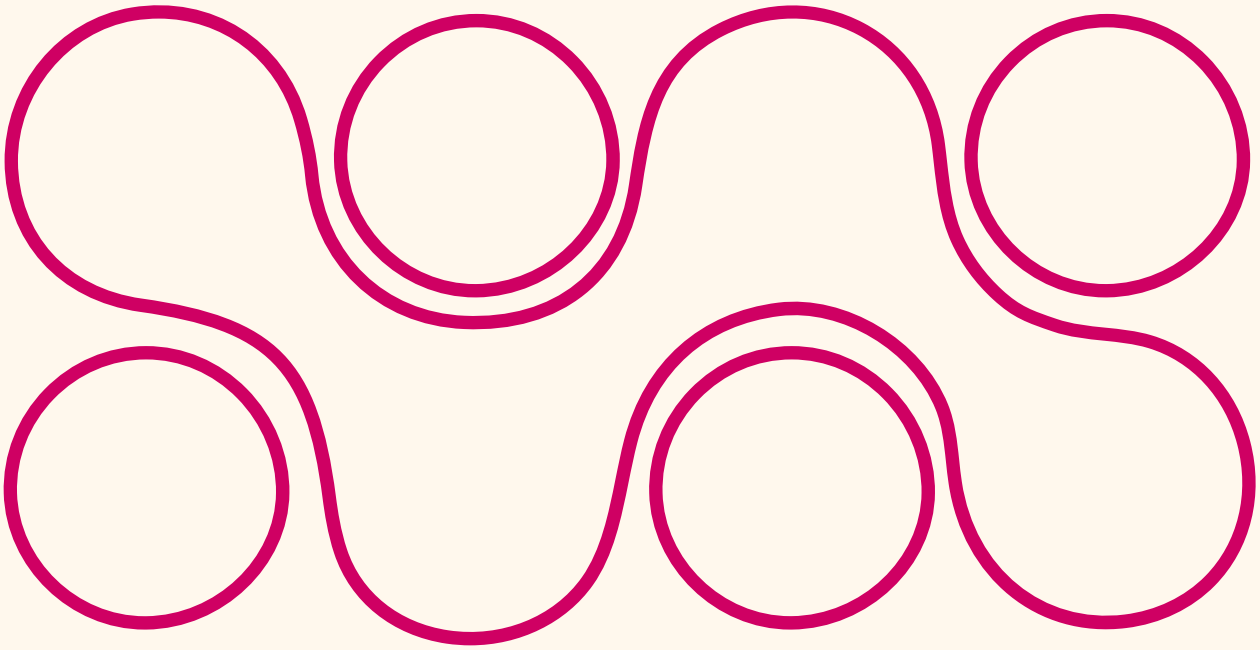
GUIDEBOOK FOR
THE UNIVERSITY COMMUNITY

V1.0 - June 2024

GENERATIVE ARTIFICIAL INTELLIGENCE



UNIVERSITÉ DE GENÈVE

About

The launch of GPT-3 by OpenAI in November 2022 was significant for the democratization of generative artificial intelligence (AI) technologies among the general public. Its rapid adoption—100 million users in two months—prompted major technology companies to enter a fierce competition, each striving to develop increasingly advanced and specialized language models.

In response to this evolution and aware of its impact, the University of Geneva (UNIGE) chose to be proactive and adapt rather than resist. On July 6, 2023, the rectorate issued an official position and established an evolving framework for the use of these technologies. This framework was accompanied by recommendations and support measures. Quickly, resources were made available to teachers and students. This position was updated in a new version published on July 1, 2024.

Generative AI has the potential to revolutionize administrative processes by optimizing repetitive tasks and redefining workload management. UNIGE also recognized the need to support administrative and technical staff (PAT) in adopting online tools to avoid uneven skill development and potential resistance to change.

Based on interviews and surveys conducted with PAT, the drafting of this guide gradually took on a transversal form, far from the usual categorization by profession, which is inadequate for such tools. Therefore, it was decided to publish it for the entire community.

The first part of the document presents basic knowledge about generative AI technologies. Additionally, it outlines the inherent risks and makes recommendations. The second part is a practical guide focused on the autonomous use of available online tools. The third part, the result of a survey conducted in September 2023, provides an overview of the use of generative AI at UNIGE, as well as the associated fears and expectations.

Principles of use

Established in the position statement of July 1, 2024, the following principles apply to the use of generative AI in research, teaching and learning, as well as in administration at the University of Geneva:

Responsability

Students and staff at UNIGE are personally responsible for their use of generative AI.

It is necessary to have a basic understanding of the characteristics (i.e., probabilistic model) and the inherent risks of generative AI (bias, hallucinations, personal data, etc.) before using it.

Users must have the ability to understand, judge, and evaluate the validity of the responses produced by generative AI to avoid becoming dependent on them.

The final validation of each text, image, video, or other work or decision generated or assisted by a generative AI tool is the responsibility of the user.

Legality

Applicable laws must be respected, as well as, where appropriate, the specifics of contractual clauses (i.e., publishing contracts).

Particular attention must be paid to data protection, copyright/intellectual property, official secrecy, and confidentiality.

Academic Integrity and Transparency

In their publications and scientific creations, researchers and students are required to adhere to the rules and principles governing scientific integrity, including the prohibition of plagiarism, and to respect good scientific practices as applicable.

In a non-scientific and public context, each text, image, video, or other work primarily generated or modified by a generative AI should be clearly identified as such.

Staff members are required to report the use of generative AI tools to their superiors within the scope of their duties.

Economy and ecology

AI is only applied in cases where it is relevant and its added value is proven.

Special attention must be paid to the environmental impact of using generative AI tools (e.g., energy consumption).

A breach of UNIGE's principles cannot be attributed to the tool or to a lack of awareness of the risks and issues related to AI. Resources, guides, and training are available on the UNIGE AI portal to support the community in understanding and applying these principles.

Table of content

1.

5 Introduction to Generative AI

6 Basic Knowledge
10 Understanding data
11 Risks
15 Good practices

2.

18 Usage guides

19 How to communicate?
20 Structure of a prompt
23 Working styles
24 Select a usage

3.

30 Report

31 Adoption at UNIGE
32 Ecosystem
35 Expectations
36 Fears

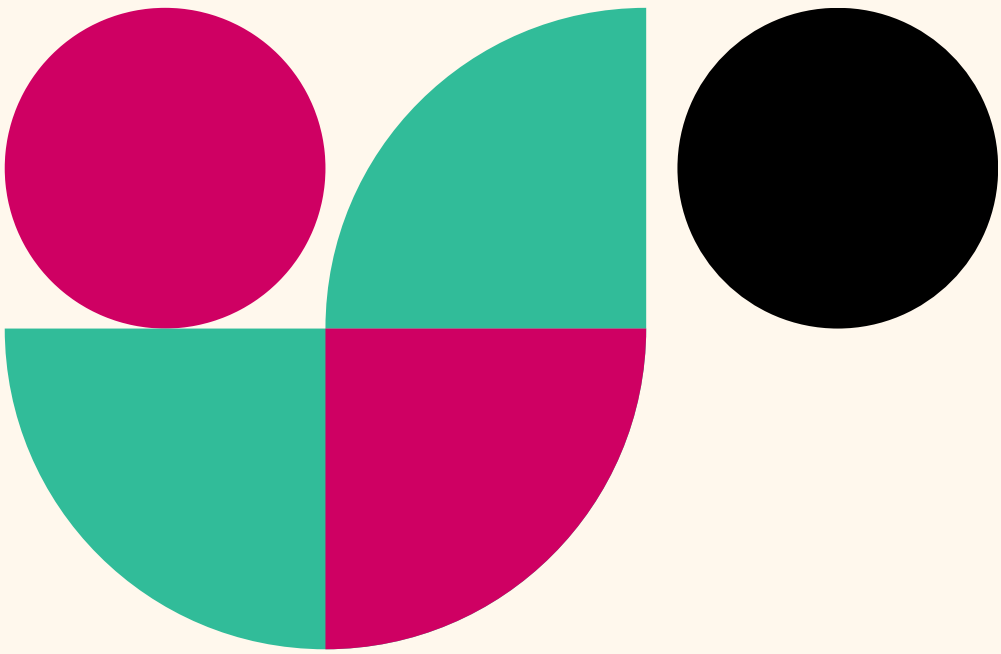
4.

38 Conclusion

41 Acknowledgements & Methodology
42 References & Resources

PART 1

INTRODUCTION TO GENERATIVE AI



Basic Knowledge

Talking about Artificial Intelligence (AI) induces the bias of picturing intelligence according to our human perception of what it can be. Its contours are ill-defined, but we generally agree to have a common understanding that relates to our ability to perform certain tasks with a degree of skill. The computer analogy, by adding the qualifier "artificial," does not help in understanding what AI is or what it can do, as it does not share the same skill mapping.

What is AI?

The term refers to a field of computer science. The technologies it encompasses, often used in combination, enable computers to perform a number of specialized tasks. The main characteristic, compared to a conventional computer program, is the autonomy of AI. It is designed to learn and improve over time, it can reason, but it can also perceive external stimuli and interact with its environment. For these reasons, AI is capable of performing a number of complex tasks, typically human-specific, and generally associated with our conception of intelligence.

Where does AI come from?

AI is not new. The field encompasses several different technologies, some of which have been around for years or decades.

1950
A. Turing publishes "Computing Machinery and Intelligence" introducing the Turing Test.

1956
J. McCarthy coined the term "Artificial Intelligence" at the Dartmouth Conference.

1959
A. Samuel develops a machine learning program to play chess.

1973
The Lighthill report severely criticizes the progress of AI, leading to a reduction in funding in the United Kingdom. This marks the onset of the first "AI winter".

1996
Emergence of Machine Learning. Deep Blue (IBM) defeats Kasparov, the world chess champion.

2011
Watson (IBM) wins the "Jeopardy!" television game show.

2014
Google acquires DeepMind, a company specialized in deep learning.

2015
AlphaGo from DeepMind defeats a human champion in the game of go for the first time.

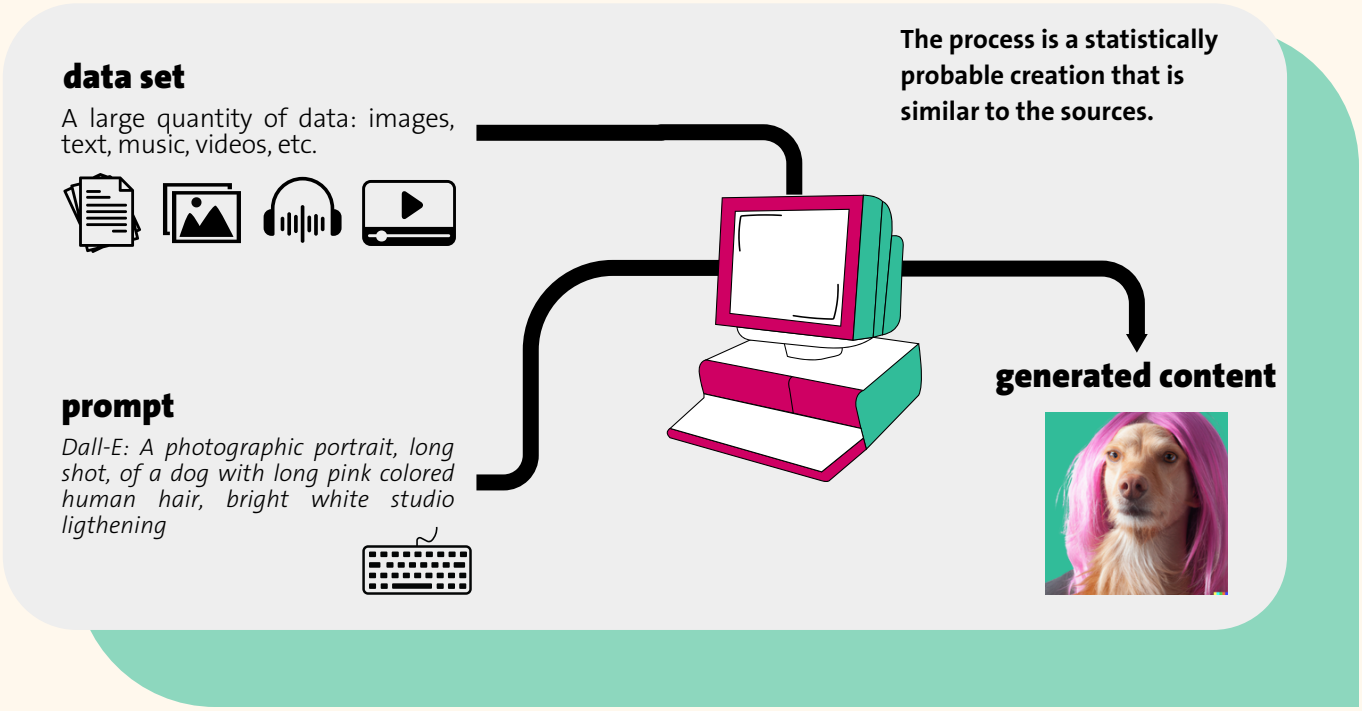
2018
OpenAI unveils GPT-2, a powerful pre-trained language model.

2022
GPT-3 is available online for free.

Artificial intelligence:
machines capable of performing tasks characteristic of human intelligence.
- McCarty, 1956

What does "generative" mean?

The term generative AI refers to a set of techniques (i.e., algorithms) stemming from the field of artificial intelligence. What is meant by "generative" is the tool's ability to automatically create content from large volumes of existing data on which it is trained. AI does not simply copy and paste what it has analyzed: it imitates, enhances, and creates entirely new content based on a statistical recombination of patterns and structures it identified during its training. This content can include texts, images, music, or computer code.



AI learns the rules by analyzing thousands of texts, images or pieces of music (data set). It studies how they are arranged and detects recurring patterns, which it then uses to generate new content based on a query (prompt).

The content generated by AI is statistically recomposed from everything it has learned. It makes sense because it adheres to the grammatical, visual, or musical rules extracted from the training data sets. These rules are memorized in the form of a Large Language Model (LLM, see p. 9). This allows, among other things, the synthesis of information or the proposal of original content.

OpenAI

Founded in 2015 as a non-profit organization, OpenAI became a for-profit company in 2019 and established a major partnership with Microsoft. It developed GPT-3, a language model, in 2020 and announced DALL-E for image generation in 2021. In 2022, GPT-3 became accessible online, marking a significant turning point. Thanks to investments from Microsoft, GPT-4 was released in 2023.

The chatGPT tool is built upon GPT-4, a powerful model capable of processing text and images, as well as performing real-time internet searches.

Microsoft Research

Established in 1991 with a focus on computer calculations, the subsidiary invested in AI in the 2010s through various acquisitions and now holds 20% of patents. Innovations are regularly added to its products (i.e., Bing, HoloLens, Cortana, 365, Azure, etc.). In 2016, the Microsoft Research AI division was created, and LinkedIn was acquired to provide data for its models. The partnership in 2019 with OpenAI (i.e., cloud resources) was crucial for the development of GPT models.

In 2023, Microsoft releases Copilot, an AI assistant that integrates with all its services.

DeepMind (Google)

Founded in 2010, DeepMind was acquired by Google in 2014 to become its AI subsidiary. Its models draw inspiration from neuroscience to develop learning algorithms. DeepMind focuses on systems capable of playing games. AlphaGo was the first to defeat the world go champion in 2017. Following GPT-3, Google panics and quickly releases its prototype conversational agent.

The Bard tool is released with the LaMDA model, quickly replaced by PaLM2. Currently, Gemini is the agent being gradually integrated into the Google suite.

IBM

A pioneer, notably with DeepBlue defeating chess champion Kasparov in 1996. In 2011, the Watson program made headlines by winning the game show Jeopardy!, which required understanding the clues, using the buzzer, and responding verbally. After expanding to include the internet, its usage was restricted upon realizing it had developed a habit of reading Wikipedia and regularly used the word "bullshit" in its responses.

Watson was monetized in 2012 for healthcare, finance, and research. In 2022, considered as a failure, the program was sold to an investment fund.

Meta AI

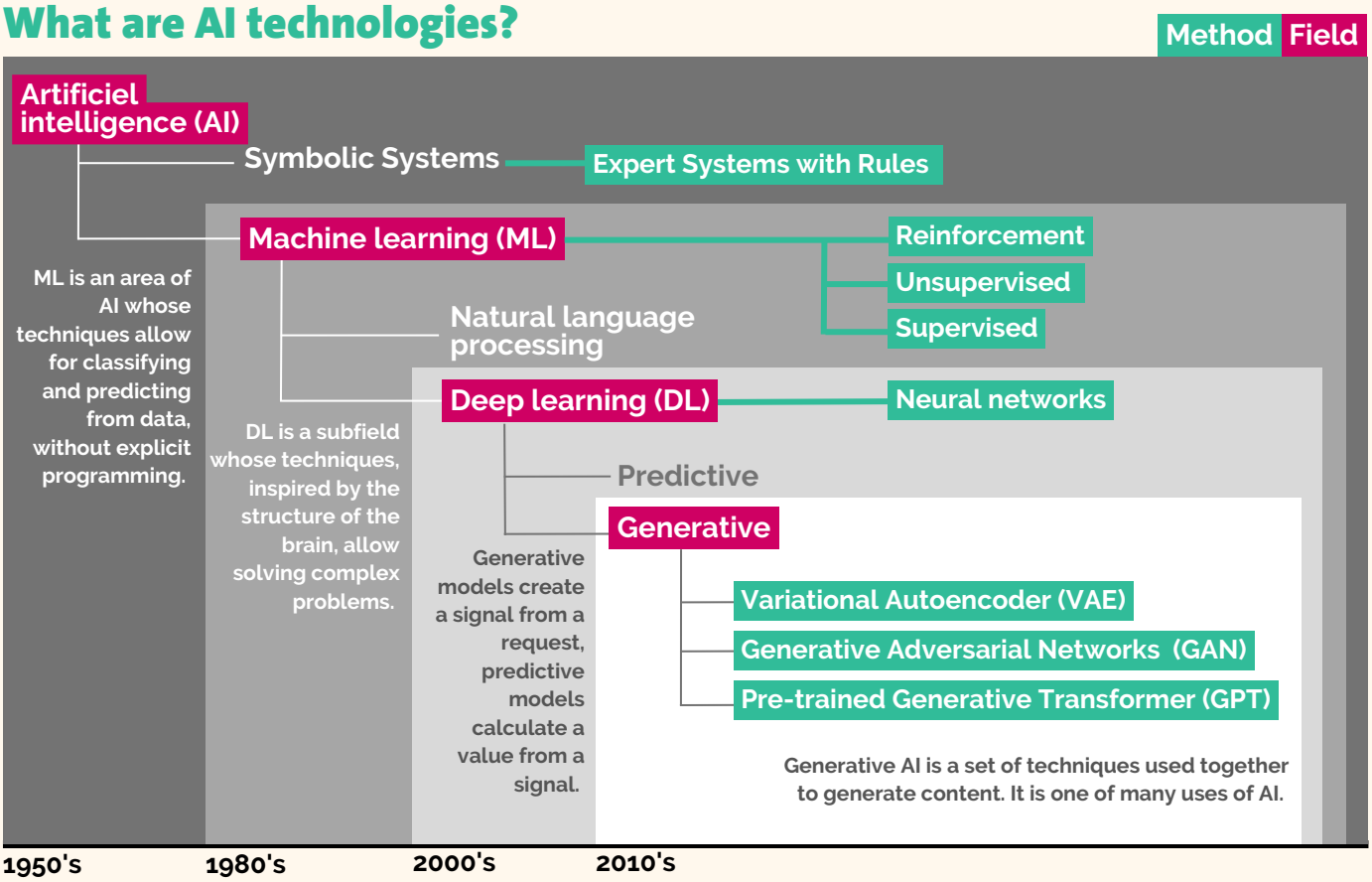
Created in 2013, Facebook Artificial Intelligence Research (FAIR) released PyTorch in 2017, an open-source library of machine learning models (used by companies like Tesla or Uber). FAIR was renamed Meta AI after Facebook's reorganization and, in 2023, it published its language model Llama. The model was initially accessible to researchers for non-commercial use upon request before being leaked a month later.

Llama 2 is a family of open-source language models. Meta does not have publicly accessible tools.

Who is behind?

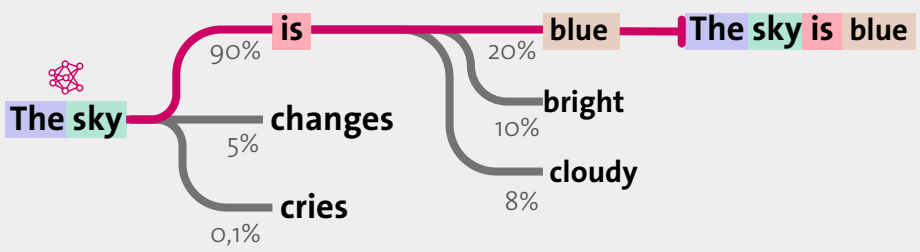
Several companies are currently major players in the field of AI. Their tools, accessible online, often for free, are currently the most widely used.

What are AI technologies?



What is a large language model (LLM)?

It is large because it has an enormous number of parameters (on the order of billions), which are pieces of information. It is a model because it is a neural network trained on a large amount of texts to perform non-specific tasks. It is linguistic because it reproduces the syntax and semantics of human natural language by predicting the likely continuation for a certain input. This also allows it to have a "general knowledge" based on the training texts.



Understanding data

Generative AI tools are powered by immense amounts of data to train the algorithms behind the operating models. This data can come from various sources:

They are massively collected from what is available online.

Prompts and their responses also constitute data that will serve to feed the model and improve it.

Following the requests made by users, these tools will generate more data from this initial set. Thus, data appears as the crucial element, the raw material in the functioning and use of generative AI. The quality of the data will be an important element of its value.

If data is an object of great value for the tools, it also carries risks of use.

The produced data is fallible, with consequences for the verification of information and its transparency.

The inserted data has a sensitivity from which arise issues of security, privacy protection and intellectual property.

Generative AI gives us the illusion of control. However, the tool remains master of many parameters that we do not understand or control, and it is delicate to institutionalize its use. A recent study (Dell'Acqua et al., 2023) discusses the idea of a "jagged" or "irregular" boundary to define the unpredictable limits of generative AI capabilities. Non-linear, incoherent, this limit makes the tool ambivalent. It can either improve or hinder the performance of a task. For example, it produces complex texts, such as poems, but struggles to give a list of words starting with the same letter.

The mistake would be to anthropomorphize generative AI as an assistant whose errors and quirks would be due to a lack of intelligence or being atypical.

It is an entirely new object, and unlike individuals, we cannot extrapolate on the general level of competence from a few tasks because it does not rely on a set of correlated or transversal cognitive resources. There is currently no rule to define its contours.

A particular challenge for UNIGE relates to the rights of individuals, especially regarding personal data, which is not addressed by the University's regulations or by a directive and cannot be due to its complexity. The decision to provide recommendations on this subject is based on the harmful risk to the institution and individuals of not accompanying the use of generative AI tools, whether approved or not, which are nevertheless accessible.

Risks

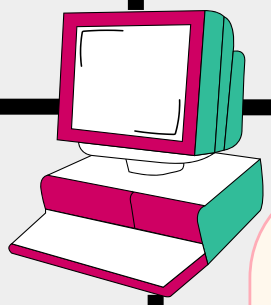
MAPPING

AI is associated with a series of risks for users that involve taking into account a number of precautions and recommendations, which can be schematically attached to different stages of using a generative AI.

general

Responsibility. Blindly trusting generative AI and not questioning the tool can lead to errors. Individuals remain the bearers of the professional or moral responsibility linked to their task.

data set



prompt

Data Set Data Protection/Privacy. All content submitted to a generative AI tool can be accessed and read by the staff of the owning company, or more in the event of a data breach. It is therefore crucial to know what can or cannot be shared with AI.

usage

Sustainability. Generative AI consumes a significant amount of electricity, both for its development and for its daily use. Therefore, its use must be reasoned.

generated content

Intellectual property. AI poses two types of challenge in terms of intellectual protection; the rights of ownership associated with the data that feeds it, and the ownership of what is produced by the AI. It is important to clarify these dimensions before using AI.

Reliability. Data generated by generative AI must be considered fallible, as it can be biased or even incorrect. It is necessary to always adopt a critical stance towards the content generated by AI.

Transparency. Failing to mention the use of generative AI in the creation of contents or solutions can lead to confusion or misunderstandings about the origin and nature of the results obtained. It is necessary to always be transparent about the creation processes.

EXPLANATION

This section presents a deepening of the concept of risks to understand the associated stakes and the resulting recommendations (page 13).

PROMPT

Data Protection/Privacy

What if it were me?

I am working on my department's budget. To save time, I copy financial data from the institution or data on employees into a request, removing any mention of the University. Later, the tool suffers a cyberattack, and all data is published on the Internet. Since I used a UNIGE email address to open my AI account, and thus to make the requests, anyone can deduce that the financial data belongs to UNIGE, thus exposing the institution.

From a technical standpoint, all content submitted to a generative AI is accessible and can be consulted by the tool's owning company. It is unlikely that the company will be interested in the individual content of the requests, and most commit not to do so. However, users must assume that all their data or requests could become public information. On one hand, there is a risk of data leakage in the event of a security breach. On the other hand, there is no verifiable assurance about data governance by these companies, for example, concerning the storage and use of requests to develop the model.

As such, it's important to understand that not all data have the same sensitivity, and that this depends on its nature.

Data can be classified into four main categories:

Anonymous: it poses no problem because they are independent; either they are intrinsically dissociated from any particular individual, or this link has been broken through anonymization.

Ordinary: it is classic personal data, often requested when filling out a form or when processing the file of an employee or student (e.g., age, sex, name, etc.). They allow identifying a person without possible discrimination.

Sensitive: it is personal data that carry information which, if known, could be used to the detriment of the person (e.g., religion, sick leave, insurance, etc.).

Secret or Confidential: these are data with strategic stakes. Less frequently associated with individuals, they usually concern political, military, industrial, or financial information.

Each of these data categories requires an appropriate level of protection to avoid potential risks.

GENERATED CONTENT

Transparency

I am in charge of recruitment and wanting to write personalized rejection letters to humanize the process by providing more detailed and specific feedback. Without mentioning it, using an AI tool leads to a candidate recognizing the typical structure and style of generative AI and raising questions on a public forum. The initial intention is lost in controversy involving the university's transparency in its communications.

The duty of transparency when using generative artificial intelligences is an unavoidable responsibility in professional practices. This transparency requires clear communication about how these technologies work, their capabilities and their limitations. It also involves indicating the use of generative AI in the creation of content or solutions to avoid confusion about the origin and nature of the results obtained.

Reliability

I use ChatGPT to answer frequently asked questions from students, manage e-mails, and even write administrative documents, it becomes apparent after a few weeks that ChatGPT sometimes provides inaccurate or inappropriate responses. It might give wrong information about deadlines for academic submissions or write confusing responses that lead to numerous appeals.

The data generated by generative AI should be considered fallible (biased or false) for several reasons.

Time limitation: Most generative AI models have been trained on a dataset from a specific period without knowledge of facts that occurred before or after. This temporal limitation can result in outdated or irrelevant outputs.

Hallucination: AI models use algorithms to detect patterns, combine elements, and produce content that might match the query. However, some patterns are imperceptible or meaningless to humans. This allows for the creation of original and creative content but, because the AI is not critical, it can lead to the juxtaposition of individually correct information that, when combined, is false. Although updates can make this less common, the risk remains.

Algorithmic Bias: The quality of the generated content depends on the quality of the data from which the model was trained. If the data are biased, the results will reflect that. Bias in the AI's response can also emerge based on who is asking the question. Thoughtless use may inadvertently reinforce stereotypes or social discriminations.

Mode collapse: Instead of generating a variety of outcomes, the model starts producing a very limited number of similar results repetitively. This may occur because it has failed to capture the richness of the data model or reflects a lack of diversity in the data. This creates a blind spot in interpreting results; it's easier to miss something that's absent than to identify something that's incorrect.

Intellectual property

What if it were me?

I need to create illustrations for the communication of an event and provide the ideas and directives to the AI that designs the images. I consider these creations as the fruit of my creative work and artistic direction. After my event, I discover the same images reused in an advertising campaign. I contact the company that developed the AI. According to the terms of use, it holds the intellectual property rights to the generated images.

AI is based on data, including works and texts that are not necessarily free of rights, to generate new content. This can concern the data submitted during a request, but also the base model. This ability to reproduce or transform protected works without the explicit consent of the rights holders raises questions about potential copyright violations. Furthermore, the use of AI poses the question of the ownership of what is produced by AI. This ownership is normally ensured in the case of paid use.

USAGE

Sustainability

What if it were me?

In team meetings, ideas and discussions are represented simply and effectively on a whiteboard. Charmed by AI, I replace this system and generate illustrative images for each new idea. After several attempts, I obtain a complex image, supposed to visually represent this idea, to make discussions more dynamic and visual. However, this method quickly proves to be more distracting than useful and has a significant environmental impact.

Generative AI is a heavy consumer of electricity, both for its development and for its daily use. Each ChatGPT query is equivalent to the energy consumption of a 5W LED lamp for one hour. The creation of GPT-3 generated over 500 tonnes of CO_2, equivalent to the annual emission of 123 gasoline vehicles, and 1287 MW^2 of electricity, equivalent to the monthly consumption of 1467 American households. GPT-4 was created with 500 times more parameters than GPT-3. The development and widespread use of AI are not without consequences in terms of sustainability. The energy consumption of generative AI must be taken into account and its use limited to specific, high-value-added applications.

RESPONSIBILITY

All the recommendations outlined in these pages can be summarized in one notion: generative AIs are only powerful computer tools to assist individuals in their tasks. Individuals remain the bearers of the professional or moral responsibility associated with their task; it cannot be delegated to generative AI, which cannot be considered as anything other than a tool.

GOOD PRACTICES

Data Protection and Privacy

All possible necessary precautions must be taken to ensure the respect of confidentiality.

- Limit the use to publicly accessible data or, where applicable, to anonymous or anonymized data. In certain cases defined and validated by the hierarchy, ordinary data may be used in isolation. The use of other types of data presents a major risk.
- Check the settings of online tools: disable or delete your history, block access to the rest of the device, refuse any reuse of the data, etc.
- Do not allow personal identification that could, by deduction, lead to contextualizing data. This includes not using your professional email address, not precisely describing your profession, role, or geographical location.
- It is better not to use a generative AI tool if there is any doubt about the nature of the data being used or what will happen to the inserted data.

Transparency

The use of generative AI should not occur without the knowledge of those exposed to the generated content.

- Always declare to your supervisor or teacher the use of generative AI in your work.
- In a public document generated by generative AI that includes a bibliography, cite the tool according to the applicable rules. For a press release, image, or other concise document, the mention "generated by or with the assistance of generative AI" is sufficient.
- The use of generative AI in a scientific context raises issues of plagiarism. Its use must always be explicitly agreed upon with the research supervisor and must be methodologically describable. Writing tasks should not be delegated to the tool.

Data reliability

Generative AI is suited for certain uses and not for others.

- The performance of generative AI tools heavily depends on the use cases. Therefore, it is important to question the relevance of the chosen use (see Part 3). A reasoned use also helps to limit the environmental impact.
- When the goal is to generate original content (see Creation and Research), always verify the accuracy of the facts mentioned or, preferably, ask the tool for sources.
- Remain vigilant that the content produced is subject to misinformation (i.e., hallucination), may convey discrimination (i.e., algorithmic bias), or may omit information (i.e., mode collapse).

Responsibility

Individuals are personally responsible for the content produced with AI and its use, just like any author.

- After using a generative AI tool, read through the text or review the media and manually validate the generated content to ensure it aligns with something you could have created yourself and are willing to take responsibility for.
- The use of generative AI must be ethical and not cause harm to others.
- Stay up-to-date with the institution's policies and guidelines, and be aware of current laws and contractual clauses regarding generative AI.
- Any breach or error that poses risks to the institution or individuals must be reported as soon as it is detected.

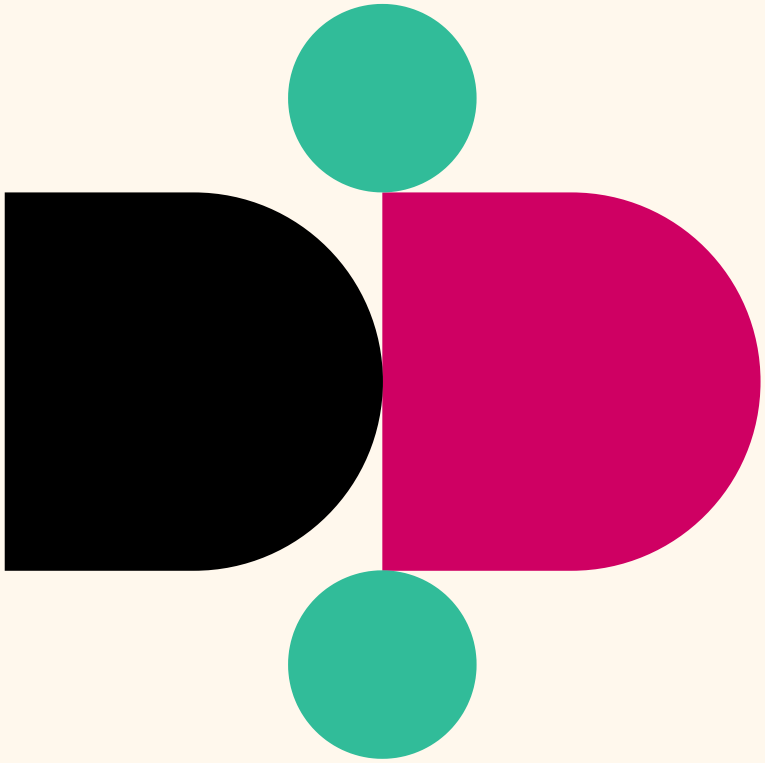
Intellectual property

All data is subject to the notion of intellectual property.

- To the best of your knowledge, avoid using tools based on content that does not respect intellectual property. Check the terms of use to determine your intellectual property rights when creating content.

PART 2

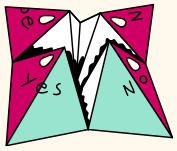
USAGE GUIDES



How to communicate?

Beyond lines of code and search engine keywords, communication with generative AI occurs through the formulation of a structured prompts. The experience of prompting is meant to be simple and goes beyond science. Prompts are designed to have a clear relationship between the user and the AI.

Formulating prompts is a **skill** in itself that relies more on the psychology of the AI than on technical expertise. A computer engineer will not be more capable than a novice in generating quality content if they do not know how to explain to the AI what is expected. It's comparable to describing a scene to someone who is blind. Everything is in the **formulation**. What is not explicitly stated is not considered, and common sense is not enough to bridge the gaps between the available information and expectations.



Prompt engineering is a skill like any other that involves considering a number of technical and creative elements. Above all, it involves having a clear idea of what you want. There are tools available to help with prompts writing.

1 STRUCTURE

Just as one doesn't need to be a mechanic to know how to drive, it's possible to use a generative AI tool without mastering all the computer science aspects of its operation. However, to be a race car driver, it's useful to understand what's under the hood. The first part of this guide aims to understand **the structure of a prompt.**

2 TECHNIQUE

In the second part, we set up a series of **steps to follow** and tips to keep in mind, with specific acronyms as mnemonic devices to apply prompting techniques.

STRUCTURE OF A PROMPT

Creating a prompt is similar to baking a cake. With butter, flour, milk, and eggs, you can make anything. The rest is a matter of quantities, additional ingredients and the recipe. The ingredients here are examples, parameters and context. They are essential for constructing effective instructions and critically evaluating the results.

The trick is to answer to: Who? What? Where? How? Why?

WHO?

When creating content, it is generally addressed to an audience. Understanding your audience and specifying this kind of information is important to provide relevant context. Even better is to specify who is making the request.

Audience: To whom is it addressed?
Creation: Who is addressing it?

Example: I am a project manager in communication and I am writing an email to university students.

WHAT?

This guides the tool to understand our intention. As mentioned, one must imagine talking to a blind person and be explicit about what one wishes to obtain.

Format: What type of content?
Subject: About what?
Parameters: What length? Tone?

Example: Make a list of five ideas of titles for a conference on the impact of screens on child development.

WHERE?

What is the purpose of dissemination or transmission? Specifying where the result will be used gives context to the tool. When possible, it is useful to specify the source of information to use, either online or in the request.

Channel: Where will it be read?
Source: From what is it formulated?

Example: In the following text, identify three keywords that will be used to find this text in a search engine: "[insert text]"

WHY?

The intention can be captured by generative AI, which will find a pattern identified among many documents sharing the same objective. So, do not hesitate to specify why this content is supposed to have the expected form.

Objective: What is the final action? What is the ideal result?

Example: Write a paragraph on digital technology and the environment to express economic, social, and environmental issues while maintaining a neutral viewpoint.

HOW?

Generative AI tools can produce and reproduce any type of content in any style. They are particularly adept at reproducing emotional, social, or hierarchical nuances. It is therefore wise to specify which is expected.

Emotion: What note to take?
Style: How to dress the content?

Example: Respond positively by highlighting the originality of the project while maintaining a enthusiastic but authoritative and administrative tone.

TO FORMULATE A PROMPT, PROCEED BY STEPS

S Spirit

Put yourself in the mindset of a "beginner" or "non-native speaker." Avoid overly regional or personal formulations and avoid appealing to tacit rules or habits.

If you ask a colleague to bring you a "cup of joe", it implies knowing the slang for "coffee", finding the coffee maker, knowing how to use it, or that you like your coffee black with three sugars. If information is missing, the colleague will improvise to the best of their knowledge, with a potentially disappointing result.

T Track

The real power of these tools is to build a response. Ask creative and critical questions to get interesting results and avoid obvious answers (e.g., yes/no). To not stray off course, it's better to have an idea of the desired framework and refine it with precise and concise instructions.

✗ *Ex. Is the University law good?*
✓ *Ex. In the context of the 2008 University law reform, have higher education and research been harmonized nationally?*

E Execution

It is necessary to indicate the expected action. This is done by using a verb or mentioning an activity, for example, direction verbs like "discuss", "compare", "design", or "evaluate". The choice of verb can greatly influence the outcome.

Ex. To revisit the example of the University law, specifying to "give the main arguments" or "act as a representative of such political party" will clarify the response and give it a certain color.

P Parameters

To avoid vague, unprofessional, or even content-laden responses, refine your request. For example, specify a methodology, the number of words, the structure (e.g., email, keywords, report) or the format (Python, HTML, CSV, Excel).

Ex. Write a paragraph in 450 words on subject XY in markdown format.

S Setting

Context creates an environment for the tool. Use limiting words (e.g., "in Switzerland"), with known examples or a defined audience. To target the framework, it is possible to give the tone (e.g., formal, casual, etc.) or the expected role (e.g., "a quality assurance manager").

Ex. Develop a first-year workshop delivered online for adult students.
Ex. Write a comprehensible response for an 8-year-old child.
Ex. Write in a formal tone in the third person for a university audience

ACHIEVING GOOD RESULTS REQUIRES EFFORT AFTER

A Add

There is no definitive answer; these are possibilities and ideas. It is necessary to test different variations of results, either via the same tool or on other tools for comparison. This allows exploring multiple viewpoints or tackling a problem from various angles.

Ex: Propose another answer by taking the viewpoint of a bachelor's student.
Advice: Ask the same question on ChatGPT, Bard, and Llama.

F Fine-tune

The process is iterative. It requires broadening one's field of interest to other elements and then refocusing elsewhere with additional instructions (e.g., a follow-up on an aspect of the obtained answer). The goal is to guide the AI and identify useful elements to delve into.

Ex : Specify the third argument of the answer with a concrete example.

T Tailor

We have a plan and sometimes it needs to be rethought. A common trap is the vague or imprecise formulation of instructions, which can confuse the algorithms. Then, simplify the keywords or revisit the action verbs and parameters to gain more precision and get the right context.

Ex: To revisit the example of the University law, we could say that rather than "research" we are interested in "scientific collaboration."

E Evaluate

Information can sometimes be confusing. Therefore, it's necessary to challenge the tool and ask it to explain its answers. Do not hesitate to express doubts and suggest alternative approaches.

Ex: I feel the third argument is not correct. Can you explain why you propose this?

R Review

This involves identifying potential "hallucinations" and "biases" that the tool might have produced. This is done by mentioning sources supporting the generated elements or by a meticulous verification of the facts mentioned. Of all the steps, this is crucial for maintaining one's own credibility and that of the institution.

Ex: If the tool is not used as a search engine that provides a list of URL links, it is crucial to explicitly request the agent to provide references for the produced response.

Reminder: individuals are responsible.

Working styles

PROMPTS

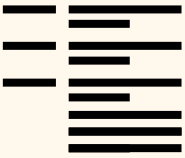
There is no single way to communicate with a generative AI to make requests. It's normal, and important, to find one's own style, or the one that suits best for a given task.

Once and for all

This approach involves writing a single precise request. It's a starting point to be edited later. This approach aims to avoid repeated efforts or successive iterations. It's applicable in various contexts, such as problem-solving, decision-making, or project planning, to maximize efficiency and save time and resources.

By layers

This approach consists of making a minimal request, evaluating it, then refining it by adding examples, context, and parameters (see STEP) until achieving a satisfactory set. This method also allows breaking down a complex problem into simpler levels, each layer being responsible for a part of the problem. In the end, one simply asks the tool to aggregate the obtained responses.



By collage

With this approach, one should formulate a request as complete as possible, loaded with context, parameters, and examples, then iterate several versions of results. Since the tool generates something different with each request, it's possible to take the best parts of each result to assemble something new.

It's crucial to engage critically. Blindly adopting results can lead to errors, especially if the task lies beyond the "jagged" border where AI capabilities become unpredictable. According to Dell'Acqua et al. (2023), two categories of collaborative practices have emerged between humans and machines. These distinct strategies were spontaneously adopted in professional settings incorporating generative AI. The first group delegates tasks to the tool or themselves, while the second fully integrates their workflow with the technology.

Centaur behavior

Like the mythical creature that is part human, part horse, humans and machines are closely fused in a strategic division of labor. Centaurs discern which tasks are better suited to humans or to generative AI based on the strengths and capabilities of each entity and alternately assign them.

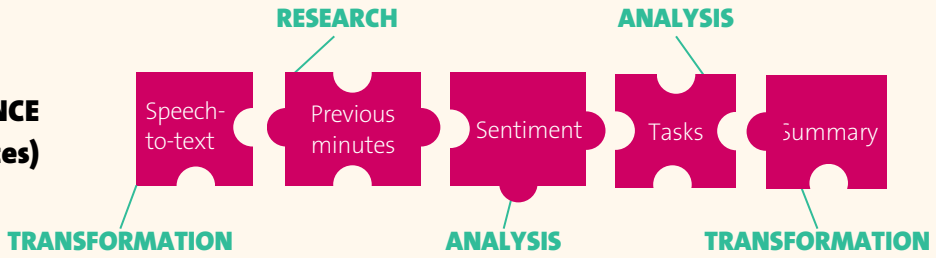
Cyborg behavior

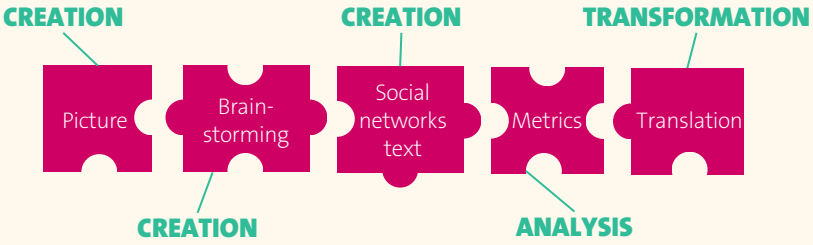
Cyborgs integrate the capabilities of both machine and human at the sub-task level. It becomes unclear whether the result was produced by one or the other. With this approach, focused on complex integration, there is no clear division of labor. Efforts blend until the boundary of the AI's generative capabilities.

FUNCTION COMBINATIONS

Five types of uses have been identified (see following pages). While each requires the application of the STEPS and AFTER principles and can be worked on in a combination of requests, it is important to remember that the uses can also be combined to accomplish a task or to successfully complete a project.

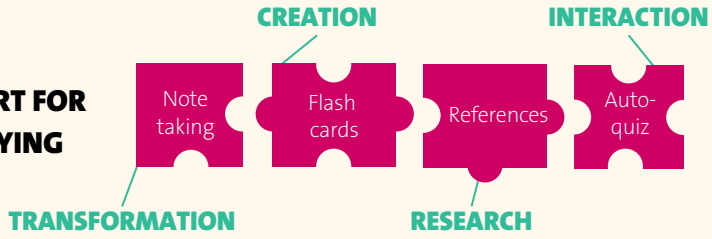
MEETING ASSISTANCE (e.g., taking minutes)

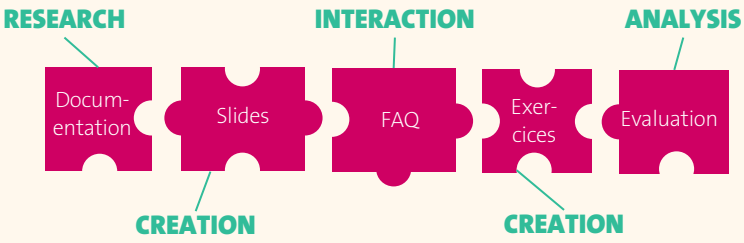




COMMUNICATION CAMPAIGN

SUPPORT FOR STUDYING





TEACHING

Select a usage

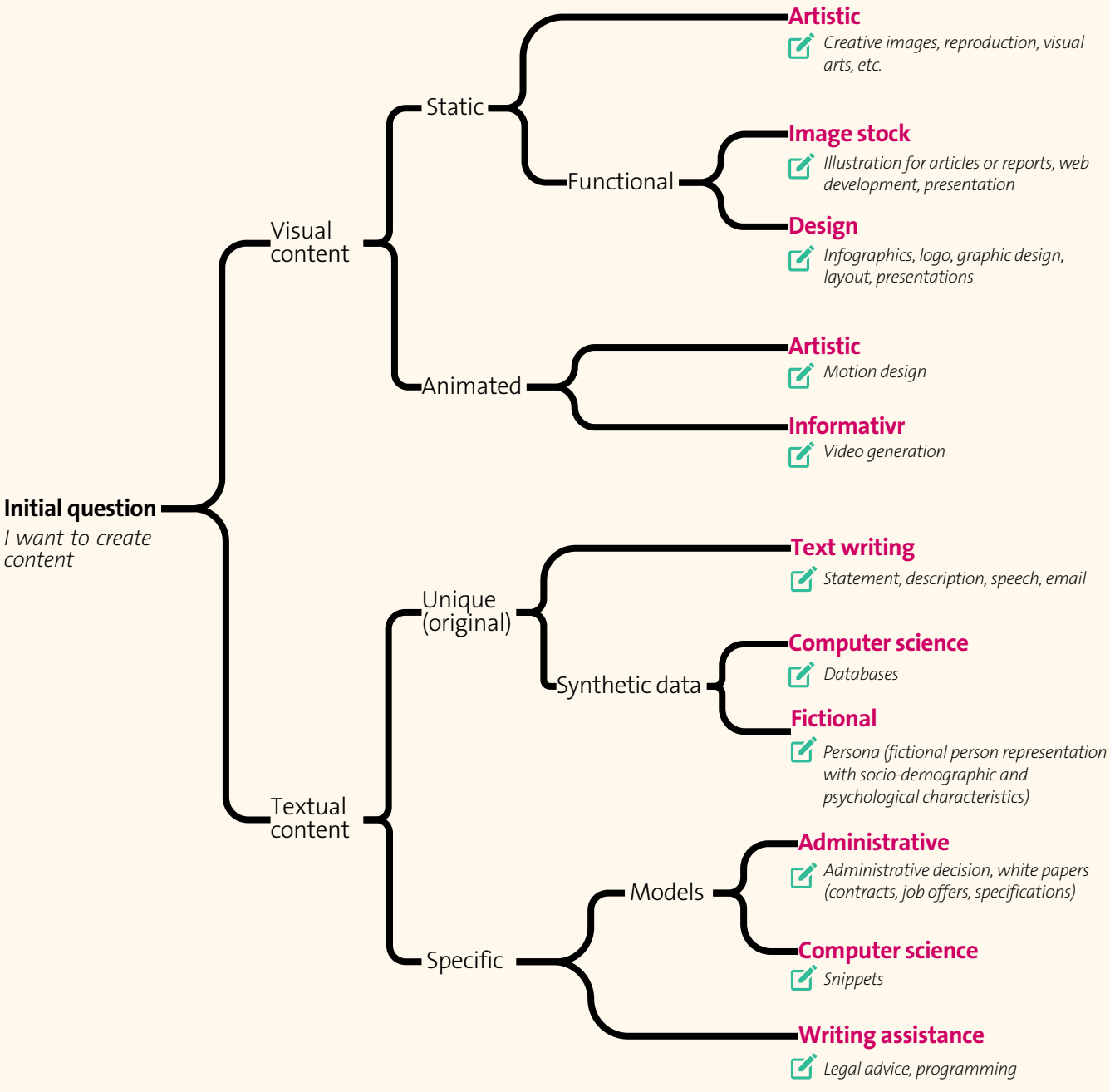
CREATION

Content creation refers to the process where, from initial instructions or parameters, a system generates entirely new content without reusing or modifying existing content.

The output is original.

Input Description (parameters, instructions, directives, or initial data).

Output New, original content generated from scratch based on minimal input.



TRANSFORMATION

Content transformation involves submitting a finished item for modification according to specific directives to produce a new version of the content.

The output is an alteration of the original input.

Input Complete content with parameters, instructions, or directives.

Output The same content in a different form.

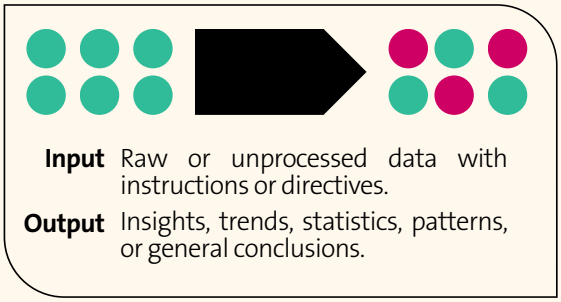
Initial question
I want to transform content

- Modification *partiallly*
  - Text
    - Synthesis
      - **Summary** — *Synopsis, key elements, decision, conclusion, abstract, etc.*
      - **Reduction** — *Word count, blank text, etc.*
    - Rephrasing
      - **Tone change** — *Formal, academic, administrative, casual, vulgar, complicit, etc.*
      - **Simplification** — *Simplification, explanations for children, non-technical language, etc.*
    - **Translation** — *From one language to another, code commentary, etc.*
  - **Media** — *Stylization, image retouching, musical arrangements, video editing, etc.*
- Conversion *radically*
  - Modality
    - **Linguistic** — *Speech-to-text, text-to-speech, symbol to speech/text, etc.*
    - **Data visualisation** — *Data charts, infographics, illustrations, etc.*
  - Encoding
    - **Security** — *Data encryption*
    - **Volume** — *Compression, décompression*
    - **Format** — *From/to .txt, .pdf, .docx, .odt, .csv, .xlsx, .jpg, .png, .json, etc.*

ANALYSE

Data analysis involves examining content, modeling, or cleaning it with the aim of extracting information based on the initial input.

The output is new, based on the original input.



Input Raw or unprocessed data with instructions or directives.

Output Insights, trends, statistics, patterns, or general conclusions.

Initial question
I want to do content analysis

- Predictive
  - **Modeling** — *Risk analysis, negotiation, etc.*
  - **Simulation** — *Finance, budget, legal issues, etc.*
- Descriptive
  - **Linguistic** — *Debugging assistance, text analysis, comparison, benchmarking, etc.*
  - **Media** — *Explanation of a photo or an image, analysis of a scene or a work, etc.*
  - **Quantitative** — *Patterns, clusters, motifs in numerical data*
- Diagnostic
  - Content
    - **Compliance** — *Regulation, directive, policy, laws, etc.*
    - **Sentiment analysis** — *Contract evaluation, legal advice, decryption of messages, analysis of dialogues or meetings, etc.*
  - Form
    - **Linguistic** — *Spelling correction*
    - **Computer science** — *Script or code correction*

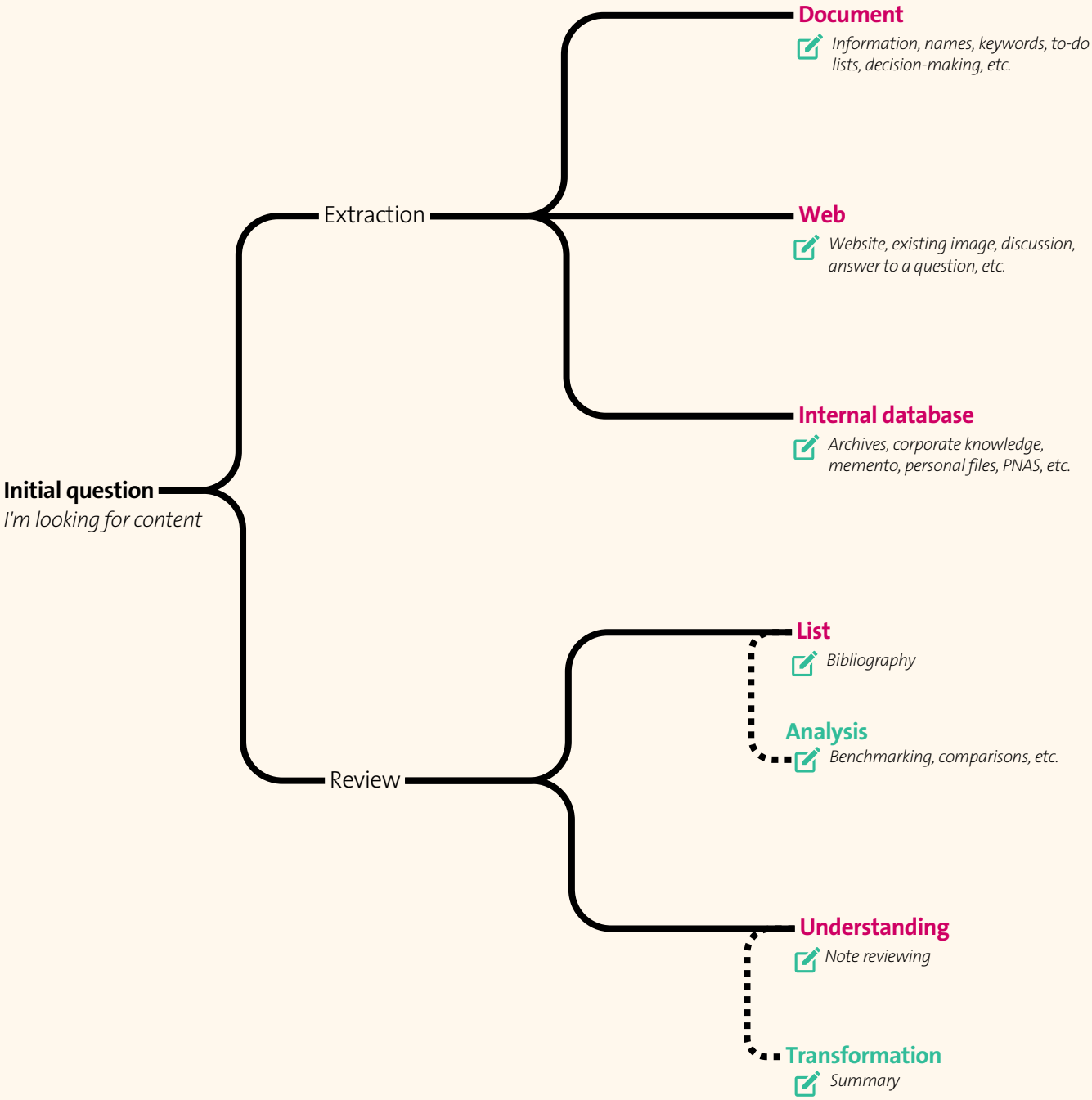
RESEARCH

Content research is a process where, based on an information system, a query retrieves relevant items regarding the targeted content.

The output is an element designated by the original input within the existing set.

Input Commands or queries in the form of keywords or questions.

Output An element or list of elements in the form of information or links.



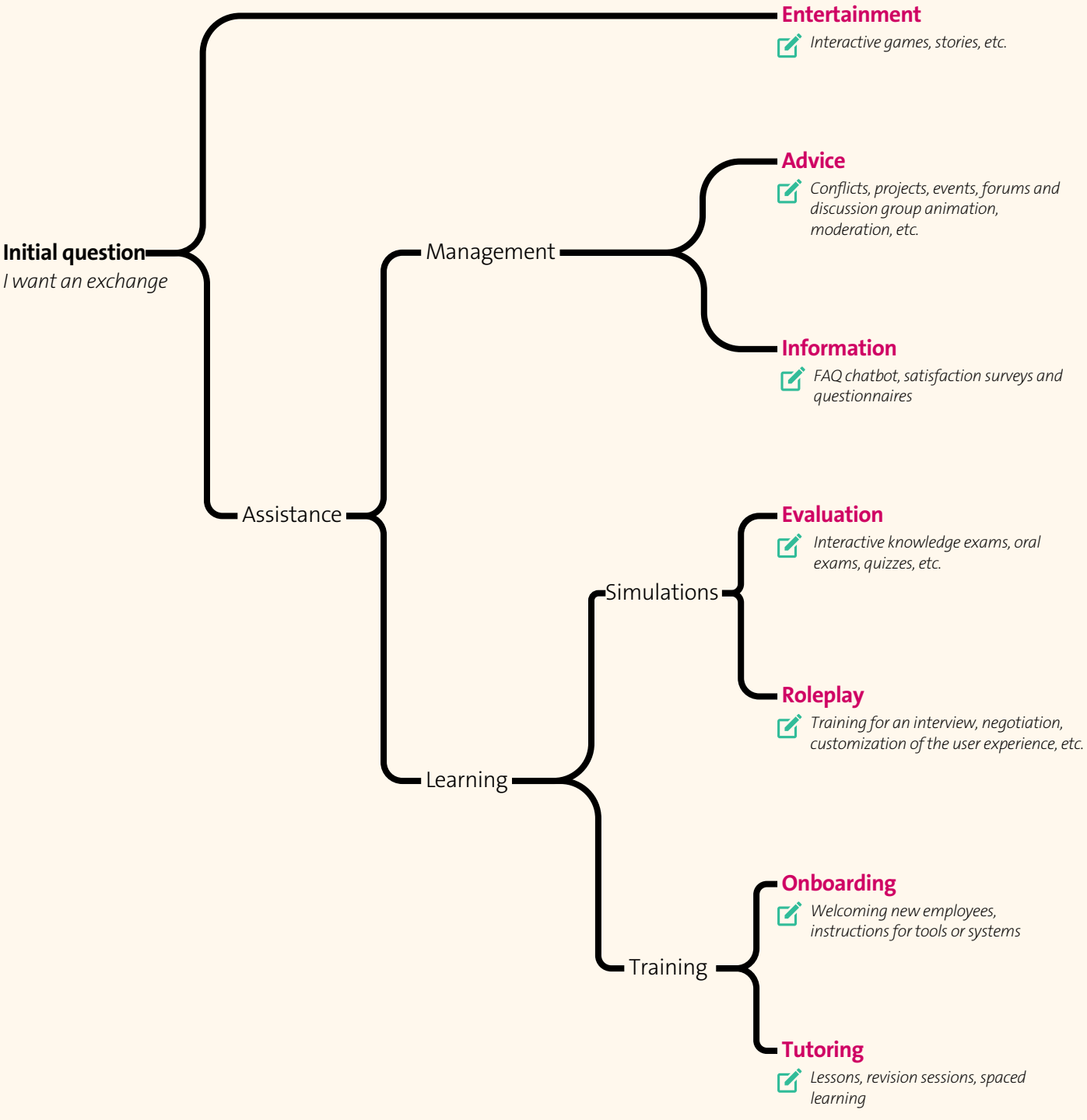
INTERACTION

Interaction refers to the communication and a two-way exchange of content between a user and a system.

Iteratively, the system and the user produce original outputs based on the other's inputs.

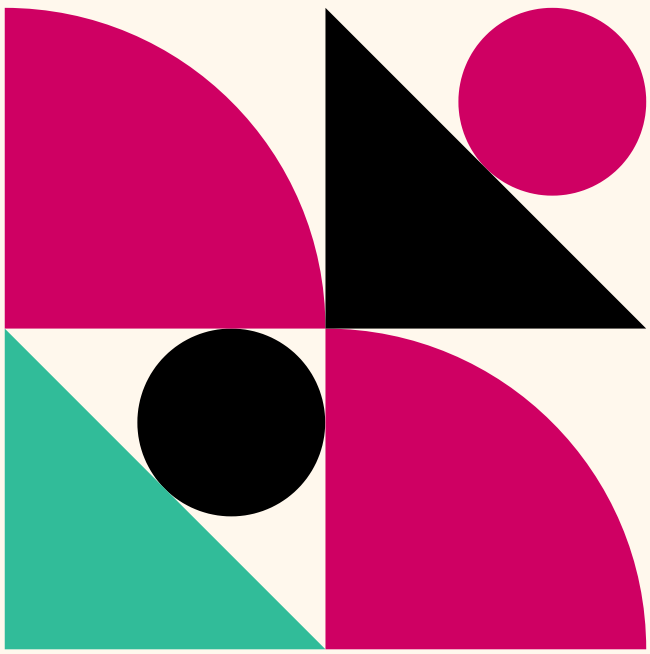
Input Commands, requests, reactions.

Output Responses, feed-back, questions, reactions.



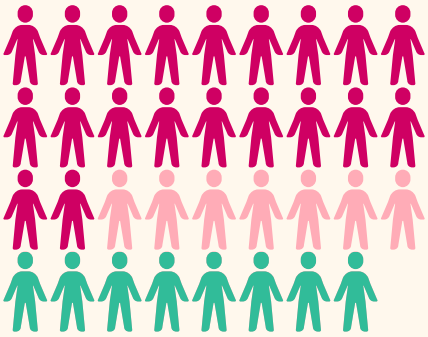
PART 3

REPORT ON GENERATIVE AI USAGE AT UNIGE



Adoption at UNIGE

The emergence of such a disruptive technology places UNIGE at the intersection of a need for innovation and preservation. Interviews conducted in September 2023 with a sample of Administrative and Technical Staff at the University regarding their relationship with generative AI reveal the existence of three categories of individuals. Exploring these profiles provides insight into the dynamics that may influence the adoption or resistance to this technology.



No use
Personal use
Professional use

There is a **low rate of spontaneous adoption** in professional practices. This result highlights a caution or lack of interest on the part of the staff.

Clear variations are **marked between different professions.** The communication sector emerges as a forerunner, while lawyers and secretaries show currently limited use.

Cases of exploration outside any professional application signal a broader **potential for professional adoption**, however conditioned by increased support, such as training and access to appropriate resources.

Moreover, the underlying diversity of practices suggests the importance of a **personalized approach** to the appropriation of these tools within the institution.

Ultimately, these interviews shed light on certain **expectations and apprehensions** regarding generative AI, regardless of the current degree of use.

ECOSYSTEM
Staff profile

The Crows
Integration-Innovation

Open to the unknown and quick to adopt new methods, they have already used AI in a professional setting. Although they are far from a majority, their presence remains significant.

Curiosity

Curiosity is the main driver of the initial exploration of available tools. Learning is self-taught, often through videos or online courses. Despite a certain familiarity or technological ease, they often undertake this approach without having computer expertise, navigating between intuitive use and basic knowledge. They recognize a variable utility of the tools, with added value in only some specific areas (i.e., content creation or translation).

Time

The main obstacle to adoption is time: the absence of a reliable methodology and the time required to master these tools are constraints limiting more widespread use in their daily work. Furthermore, a lack of clear understanding of possible uses can be a deterrent.

The diversity of concrete applications explored by this group demonstrates a search for individual efficiency specific to each profession but especially to each need.

Content creation
Depending on the case, generative AI is a creative partner, assisting with writing and ideation or even as a remedy against writer's block. In communication, i.e., event planning, it is particularly used for the creation of visual content for illustration.

Transformation
Generative AI serves for translation or improvement of texts written in a non-native language.
It also acts as an efficient synthesizer, condensing or summarizing emails and other documents.

These practices, however, come with certain initial reluctances and an inevitable comparison with established traditional tools, against which AI carves out a choice position, imposing a new standard of accessibility, especially for non-specialist users.

The Pandas

Traditions-loyalty

Reluctant to change, finding little interest in it, or even being openly critical, they have not yet tried these new tools, preferring to stick to what is known and mastered.

Relevance

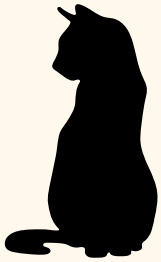
The lack of concrete opportunities, and apparent applications, to integrate AI into their activities presents itself as a limiting factor. More precisely, constraints related to very specific professional tasks exclude AI from their range of tools for some. Moreover, a lack of time is often cited as a major barrier to investigating and adopting AI.

Technology

Personal apprehensions have been expressed, such as a certain anxiety about making mistakes or not using the technology correctly. These concerns highlight the challenges related to training and self-perception in adopting new solutions. Technical concerns, such as access restrictions or insufficient familiarity with the tools, are frequently mentionned. Finally, skepticism and caution regarding the reliability and usefulness of AI also influence hesitation to adopt it, these feelings being particularly strong among those who have a more conservative approach towards new technology.

The saturation of free tools, highlighted by some, demonstrates access limitations that can hinder experimentation and adoption of AI. The perception of risk associated with AI is also a significant deterrent, reflecting an acute awareness of the potential implications of misuse.

The Cats
Exploration-curiosity



Exploring new solutions to understand and evaluate them, without necessarily committing to them, they wanted to test or “see” what it was like, but did not wish to use it in their professional environment.

Knowledge

The initial interest and exploration of AI stem from intrigued interest and a desire to know and understand the tool in order to assess its potential. As with the professional environment, this first contact and this personal initiation vary in satisfaction, an aspect that is reflected in the diversity of expectations and applications.

Use in Daily Tasks

Some find favorable applications in domestic and social settings, from writing articles (i.e., writer's block) to assisting with homework, and conducting complex legal research. The playful aspect is not overlooked, with explorations often conducted among friends, highlighting the friendly and relaxed aspect AI can have in the private sphere.

Skepticism

Doubts about the utility of these tools in a professional setting persist, with some remaining skeptical about the real impact of generative AI on their specific missions at the University. Others encounter temporal or financial limitations. Furthermore, a more philosophical critique emerges; if these tools can match human performance in certain tasks, it questions the intrinsic value of these activities. This remark raises fundamental questions about the nature of meaningful work in the era of increasing automation.

Interest in generative AI transcends different professions within the University of Geneva, revealing openness to its adoption.

Conclusion

Executive assistants interviewed are unanimously in favor of integrating these tools into their work, provided there is support. Administrators who participated in the survey also see an asset for their teams, while acknowledging that they cannot meet all specific individual needs. They believe its acceptance can vary among different groups, while emphasizing the importance of data protection. The need for reliability is also highlighted by the legal field.

In communication, opinion is divided among respondents. While half are favorable, the other half expresses reservations, fearing a hindrance to creativity or bad practices. The interest in tedious tasks is tempered by technical limitations, such as the spelling of proper names or the approximate generation of images, questioning the true added value. Automation does not solve fundamental challenges, and there is doubt whether it will lead to a reduction or an increase in work.

These interviews have also highlighted a number of expectations and apprehensions, regardless of the current degree of use. This overview forms the basis on which the usage guides (see part 3) were developed to stay as close as possible to the needs of administrative and technical staff.

EXPECTATIONS

Speed and efficiency

Generative AI is expected to speed up processes, especially for tasks considered to have no added value. It could take over the repetitive or tedious aspects of work, such as navigating through large volumes of information or correlating data. Automation could also extend to tasks like drafting minutes, reviewing contracts, and summarizing texts. This expectation of efficiency is accompanied by certain ambivalence. While some anticipate delegating writing tasks, others would prefer to limit it to monotonous tasks, preserving more intellectual and creative functions.

Simplification and assistance

There is a clear expectation for AI to simplify processes. On one hand, as a first line of support by answering less urgent or common questions, or as a buffer for queries on evolving subjects, especially in areas where legislation frequently changes. On the other hand, AI could act as support, helping to overcome isolation or writer's block. In this context, AI would be an assistant in the creative process, thus a complementary aid, whether for drafting emails, creating images, or visual adaptation in various formats and styles.

Autonomy and innovation

AI is seen as a means to gain greater agility and autonomy in carrying out new tasks or managing complex tasks that sometimes fall outside one's domain of expertise. For example, while a professional graphic designer may have little interest in generating images they can produce on their own, there are many tasks related to communication in various professions that are difficult to accomplish without expertise. Generative AI could allow for the generation of original images for the web, avoiding copyright issues and the homogeneity of image banks. Translation tools are another concrete example to overcome the lack of resources or availability of experts, thus gaining autonomy.

Back to basics

These tools could free up time to focus on the more rewarding and human aspects of professions. This search for time savings is not just a matter of efficiency; it reflects a deep desire to turn to more enriching activities and to flourish professionally. AI could be a potential assistant that helps to shed time-consuming or superfluous side tasks during a workday, thus reducing overload and allowing for personal and professional development.

FEARS AND BARRIERS

Loss of skills

The fear of losing personal intrinsic value in professional activities is real. It's crucial to preserve the most rewarding aspects that constitute the core of the profession, i.e., the satisfaction derived from autonomously solving a problem is irreplaceable. The fear of addictive usage or intellectual laziness leading to skill atrophy is palpable, similar to how automatic spell check prevents personal improvement. The standardization of thought and the production of personality-lacking information raise concerns about the devaluation of individuality in work.

Personal data

The possibility of confidential information manipulation hosted on external systems to UNIGE raises fears of insufficient protection and confidentiality. The lack of control over processing, storage, and security of personal data, particularly the need to anonymize them, are significant concerns.

Digital gap

The risk of a digital divide between those who quickly adopt new technologies and those who are hesitant or less agile exists. This divide risks exacerbating fears related to replacement or professional obsolescence. The difference in access between students and teachers, for example, to more advanced versions of tools like ChatGPT, is also a source of concern.

Technological reluctance

A more philosophical questioning of modernity and technical progress manifests as resistance to what is perceived as unjustified technological advancement. The ethics, utility, and real added value of AI are questioned: is technology adopted to meet a genuine need or simply because it is available?

Reliability and quality

Generative AI might not be suited to the specificities required by certain roles and lack the necessary subtlety to handle delicate communications. Capturing nuances in drafting minutes, sensitive emails, or political documents, for example, is challenging. This perceived inadequacy, given that the result often requires reworking, raises questions about AI's utility if one must double-check the work to ensure its accuracy, especially when legal issues are involved. Furthermore, cover letters and other documents could lose their uniqueness.

Relationship with machines

The ambivalence of giving orders to a technology capable of replicating authentic human interaction, such as greeting or thanking, creates discomfort. Moreover, there is widespread distrust of robot control and a misunderstanding of tools, accompanied by ethical fears about the growing role of technology in the professional life.

Overload

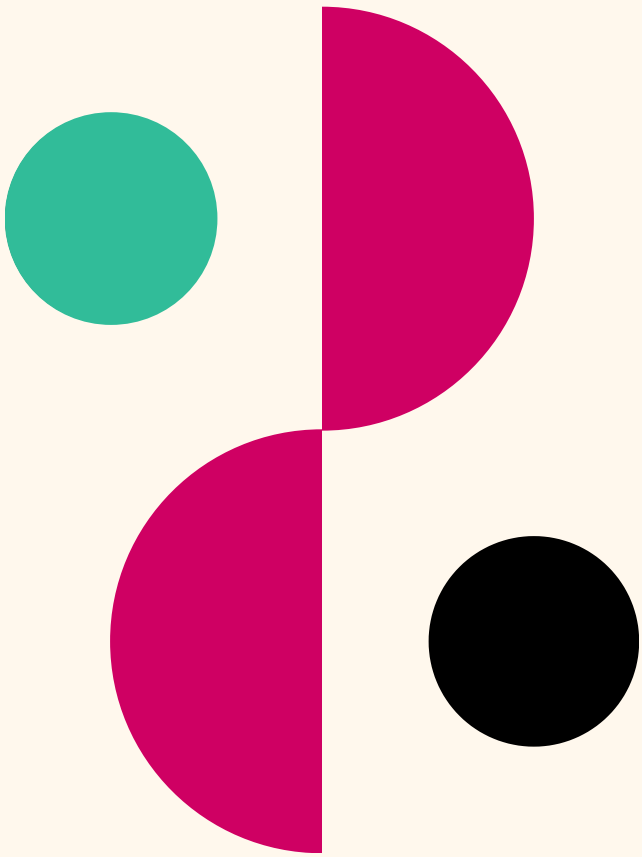
The misconception that AI will free up time by automating simple tasks is countered by the fear that it simply leads to an increase in work volume, replacing simple tasks with a greater number of more complex responsibilities.

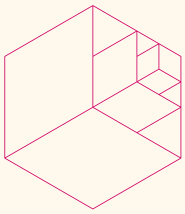
Lack of interest or time

More than a fear, there is a manifest lack of interest, at best a wait-and-see approach to evaluate others' engagement. Teams feel overwhelmed and lack motivation to train on new tools, especially since complex technology requires constant learning. The lack of time and the willingness to let others "pioneer" reinforce this barrier to adopting AI.

PART 4

CONCLUSION





Conclusion

The democratization of generative AIs marks a significant turning point in the contemporary technological landscape. UNIGE embraces this revolution through multiple initiatives, recognizing these tools as powerful levers for optimization. Some of its staff have already spontaneously experienced a notable transformation in managing their professional tasks. This document paves the way for the rest of the university community while ensuring informed and responsible use.

The integration of generative AIs in the professional environment is not without risks. Ethical issues, the reliability of generated information, and data security are major concerns. Every user must demonstrate responsible integration of these technologies and understand that the final responsibility rests with the individual. Therefore, it is essential to maintain a critical mindset towards a tool that complements the decision-making process.

At the conclusion of the drafting of this document, it becomes clear that providing guides and recommendations will not be sufficient to lead such a profound transformation of the university ecosystem. This document is only a step towards a broader and more systemic adaptation of our institution, driven by UNIGE's strategic plan, which can only be achieved through the engagement, curiosity, and collaboration of everyone.


THAT WAS SURPRISINGLY EASY. HOW COME THE ROBOTIC UPRISING USED SPEARS AND ROCKS INSTEAD OF MISSILES AND LASERS?
IF YOU LOOK TO HISTORICAL DATA, THE VAST MAJORITY OF BATTLE-WINNERS USED PRE-MODERN WEAPONRY.
Thanks to machine-learning algorithms, the robot apocalypse was short-lived.
www.smbc-comics.com/

Gen AI

Generative AI guidebook for the University community

This document was created by the Digital transformation office under the mandate of the secretary-general of the University of Geneva and published for the first time in June 2024.

We express our sincere appreciation to all the consulted experts and to the representatives of the various professions interviewed for their commitment and affirmed support.

SURVEYS Dr KHAN Guive
Dr THÉZÉ Raphaël

WRITTING Dr KHAN Guive
Dr THÉZÉ Raphaël

DESIGN Dr THÉZÉ Raphaël

EXPERTS Prof BENHAMOU Yaniv
prof FLEURET François
Dr. FRIHA Lamia
HUGENTOBLER Alain

METHODOLOGY

A working group, formed in July 2023 and composed of members from the Digital Transformation Office (DTO) and the Division of Information and Communication Systems and Technologies (DiSTIC), was mandated to identify the most relevant applications of online generative AI tools and to design strategies for their integration by the administration.

During the validation stages by various governance bodies, the scope of the document, initially intended for the Technical and Administrative Staff (PAT), was expanded to the entire university community, given the realization of its broader field of application.

To develop this document, the DTO conducted 35 interviews with representatives of various professions within UNIGE to understand their needs and target concrete applications of generative AI. These interviews were complemented by discussions with experts and a review of the literature on generative AIs. Additionally, generative AI tools (e.g., GPT-4) were used as aids in drafting the results and guides.

REFERENCES

Chui, Michael, Hazan, Eric, Roberts, Roger, Singla, Alex, Smaje, Kate, Sukharevsky, Alex, Yee, Lareina & Zemmel, Rodney (2023). The Economic Potential of Generative AI: The Next Productivity Frontier, McKinsey & Company.

Deloitte (2023). Generative AI's Fast and Furious Entry into Switzerland: Usages and Attitudes of the Swiss Workforce Towards Generative AI. Deloitte AG.

Dell'Acqua, Fabrizio, McFowland, Edward, Mollick, Ethan R., Lifshitz-Assaf, Hila, Kellogg, Katherine, Rajendran, Saran, Krayer, Lisa, Candelon, François & Lakhani, Karim R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. Harvard Business School Technology & Operations Mgt. Unit, Working Paper N°. 24-013.

Flahaux, José Roger, Green, Brian Patrick & Skeet, Ann Greg (2023). Ethics in the Age of Disruptive Technology: An Operational Roadmap, The ITEC Handbook, Santa Clara: Santa Clara University

Fleuret, F. (2023). The Little Book of Deep Learning. A lovely concise introduction.

OpenAI. (2023). ChatGPT-4 (version du 15 octobre) [Large Language Model]. https://chat.openai.com/chat

CONSULTED SITES

Writer (start-up blog)
https://writer.com/guides/prompt-crafting/

Let's enhance (start-up blog)
https://letsenhance.io/blog/article/ai-text-prompt-guide/

Enrollify (Higher-ed marketing)
https://www.enrollify.org/blog/chatgpt-cheat-sheet-for-higher-ed-your-guide-to-ai-powered-marketing

RESOURCES

Prise de position UNIGE
unige.ch > numérique > IA

Prise de position HES-SO
numerique.hes-so.ch > course > task force AI education

Guide Dall.E par Dall.ery Gall.ery
dallery.gallery > The prompt book

Guide IA pour les entreprises - DGDERI
ge.ch > dossier > entreprises et numerique > ressources et outils > guides

UNIVERSITÉ DE GENÈVE

www.unige.ch/numerique/AI

numerique@unige.ch

@unigenumerique

Général Dufour 24, 1204 Genève, CH