

FaktaBaari **EDU**

Digital Skills Needed to Identify
Disinformation

21academy
15.11.2023
Kari Kivinen

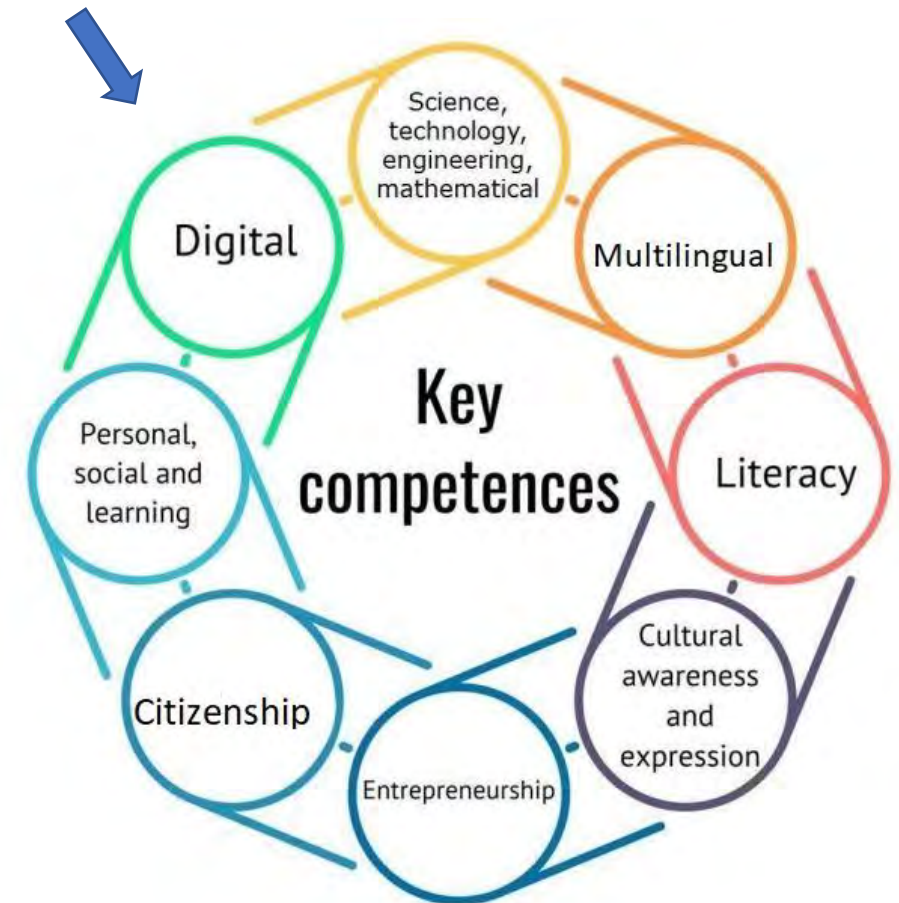


Key competences for lifelong learning - Digital competence

“Digital competence involves the confident, critical and responsible use of, and engagement with, digital technologies for learning, at work, and for participation in society.

It includes information and data literacy, communication and collaboration, media literacy, digital content creation (including programming), safety (including digital well-being and competences related to cybersecurity), intellectual property related questions, problem solving and critical thinking.”

[\(Council Recommendation on Key Competences for Life-long Learning,](#)



People with basic or above basic digital skills, 2019

(%, share of people aged 16-74; during the 12 months preceding the survey, by degree of urbanisation)



Note: ranked on cities

(1) Low reliability

(2) Rural areas: low reliability

(3) 2017 data instead of 2019

What does it mean to be digitally competent

The EU has set ambitious targets for at least 80% of the population to have basic digital skills by 2030.

- DigComp 2.2. provides a common understanding of which are the key areas of digital competence.
- Digital Literacy competences are considered as basic civic skills

<https://publications.jrc.ec.europa.eu/repository/handle/JRC128415>



EU's Digital competence framework for citizens – DigComp 2.2.



DigComp 2.2. – Information and data literacy

	EXAMPLES OF KNOWLEDGE, SKILLS AND ATTITUDES
Knowledge	<p>16. Aware that online environments contain all types of information and content including misinformation and disinformation, and even if a topic is widely reported it does not necessarily mean it is accurate.</p> <p>17. Understands the difference between disinformation (false information with the intent to deceive people) and misinformation (false information regardless of intent to deceive or mislead people).</p>
Skills	<p>24. Knows how to differentiate sponsored content from other content online (e.g. recognising advertisements and marketing messages on social media or search engines) even if it is not marked as sponsored.</p> <p>25. Knows how to analyse and critically evaluate search results and social media activity streams, to identify their origins, to distinguish fact-reporting from opinion, and to determine whether outputs are truthful or have other limitations (e.g. economic, political, religious interests).</p>
Attitudes	<p>29. Willing to fact-check a piece of information and assess its accuracy, reliability and authority, while preferring primary sources over secondary sources of information where possible.</p>



DIMENSION 1 • COMPETENCE AREA

1. INFORMATION AND DATA LITERACY

DIMENSION 2 • COMPETENCE

1.2 EVALUATING DATA, INFORMATION AND DIGITAL CONTENT

To analyse, compare and critically evaluate the credibility and reliability of sources of data, information and digital content. To analyse, interpret and critically evaluate the data, information and digital content.

Information disorders

Misinformation

Verifiably false information that is spread without the intention to mislead, and often shared because the user believes it to be true.



Disinformation

Verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public. It can cause public harm.



Malinformation

Factually correct information that is used harmfully

Mistake

Lie

Gossip

Eurobarometer - Malta

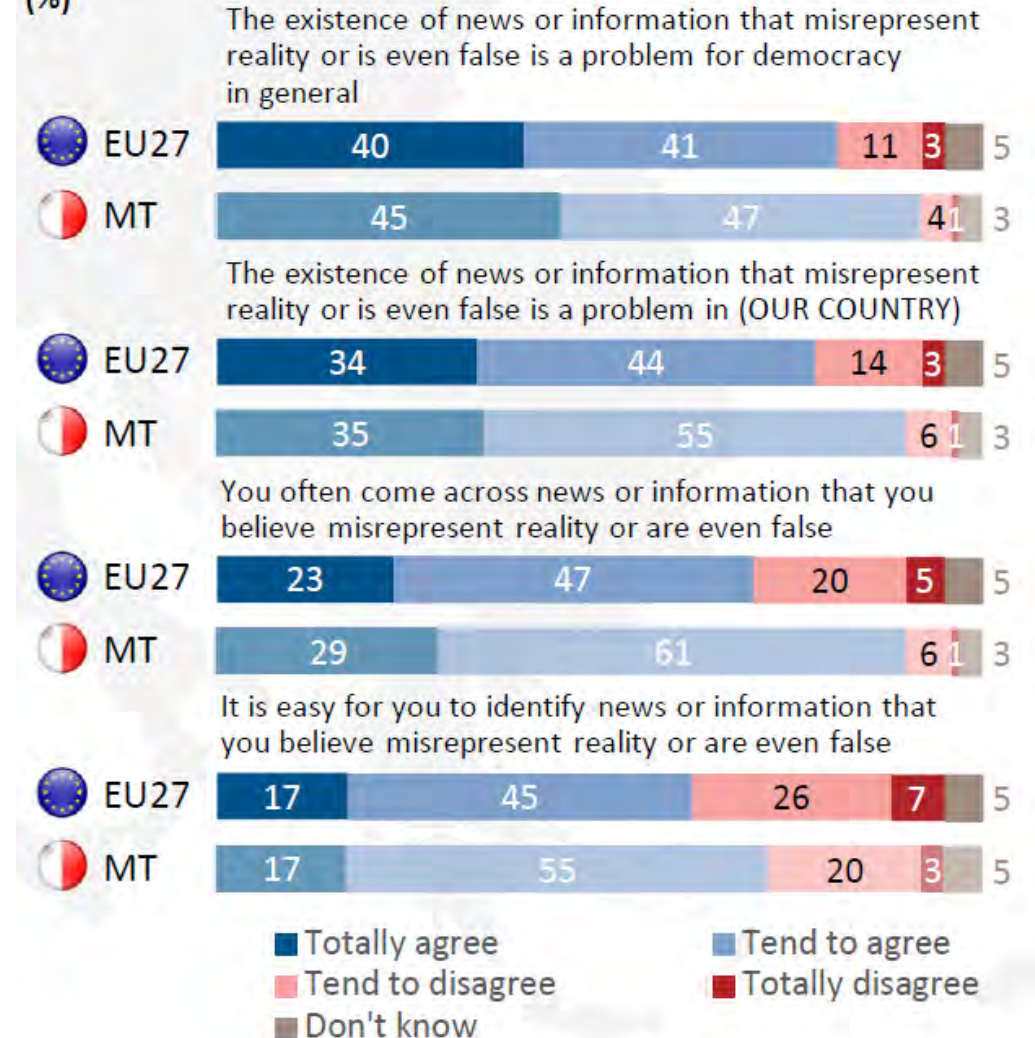
According to Eurobarometer (2021/2022)

- ❑ Mis/disinformation is a problem for democracy 92 %
- ❑ Mis/disinformation is a problem in our country 90 %
- ❑ Coming across mis/disinformation 90%
- ❑ It easy to identify mis/disinformation 72 %

<https://europa.eu/eurobarometer/surveys/detail/2553>

QD8 Do you totally agree, tend to agree, tend to disagree or totally disagree with each of the following statements.

(%)



Fact-checking service in Malta

- Fact-checking is the process of verifying if information is true or false.
- It can take place with any type of media (and for instance for both text and visual images).
- There are many claims which cannot be fact-checked.
- Fact-checkers use often image and video verification tools to verify the authenticity of the photos or videos.

The screenshot shows the Times of Malta website's 'Fact-check' section. The page is titled 'timesofmalta.com/articles/listing/fact-check' and features a navigation menu with categories like 'Latest', 'National', 'World', 'Opinion', 'Fact-check', 'X2', 'Sport', and 'Motoring'. The 'Fact-check' category is highlighted. Below the navigation, there are three fact-check articles:

- Fact-check: Is the Ġgantija development outside UNESCO's buffer zone?**
The project's applicants say that it is, but objectors argue otherwise
2 hours ago | Neville Borg | 7
- Fact check: Will running hospitals cost as much as Steward received?**
Robert Abela said that running the hospitals is 'certainly not cheaper' than what government was paying Steward
November 9, 2023 | Neville Borg | 49
- Fact-check: Scam posts mimic Times of Malta interview with Robert Abela**
This is the latest in a line of scams showing in-depth knowledge of local issues and personalities
November 2, 2023 | Neville Borg | 8

Each article includes a thumbnail image with a red 'FALSE' or 'MOSTLY FALSE' label and a caption describing the fact-check. For example, the first article's thumbnail shows a street scene with the caption 'Fact-checking whether the Ġgantija project is outside a UNESCO buffer zone'.

Online vs. offline environments

- In the online environment, the amount of information available is breathtaking and it is possible to diffuse any information effortlessly to vast audiences in no time.
- Contents can be changed, removed and added all the time.
- The results of the search engines and recommender systems are individualised and unpredictable.

Online environments are often designed to:

- maximise commercial interests,
- capture and sustain users' attention,
- monetise user data, and
- predict and influence future behavior

Kozyreva et al (2020)

Online reading skills & strategies

- ❑ Traditional reading skills should be complemented by new online assessment strategies and online literacy skills to tackle disinformation!

Effective methods proven to tackle disinformation:

- prebunking (anticipation),
- debunking (correction),
- strategic ignorance
- lateral reading
- civic online reasoning
- etc

Debunking & pre-bunking

What is debunking?

Debunking happens after the fact, so after false information has appeared.

The aim is to correct false information and to prevent others from believing what is verifiably false information.

Those reading or seeing the information 'see through' what is being presented as fact and/or truth.

Fact-checking strategies can be used to debunk misinformation and disinformation.

What is pre-bunking?

Pre-bunking is a process where people are warned in advance that they are about to be the target of false information.

Pre-bunking can be taught to students by providing them with factual and some in-depth information on a particular subject beforehand, and then introducing the existing disinformation about the same subject.

They can also be told in advance what kinds of disinformation they can expect.

Strategic ignorance

- ❑ When using powerful search engines, we sometimes get millions of hits.
- ❑ How to select information which is useful, truthful and which meets our initial information need? In this process we need human critical thinking to evaluate the value of the content algorithms are proposing for us – and we have put aside and ignore, most of the hits.
- ❑ Advertisers, corporations, lobbyists, clickbait sites, conspiracy theorists, hate groups, and propaganda-fuelling governments work overtime to hijack our online attention.
- ❑ Often the wisest thing to do is to preserve attention by practicing strategic ignoring. Under conditions of limited attention, the most crucial decision to make is where to allocate it.

- So, we must develop skills to ignore great amounts of non-important information.
- We should embrace strategic ignoring to avoid disinformation and to preserve our limited amount of attention on content which is really worth reading.

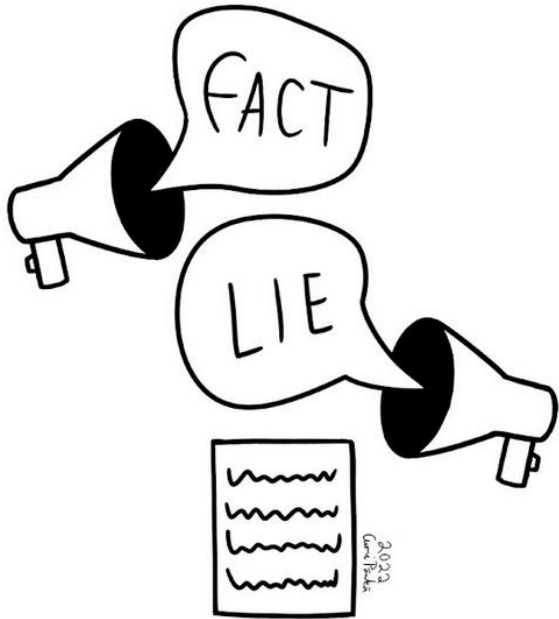
Lateral reading

- ❑ Fact-checkers use lateral reading to search information.
- ❑ Instead of opening a suggested link and exploring each hit in depth, they laterally check the background of the information.
- ❑ This is done by opening new windows in the browser, looking for answers to at least for the three key questions!

Lateral reading

The reader checks the background of the online information (author's credibility, facts, statistics, sources, etc.) on various sites and sources before reading the text at hand.

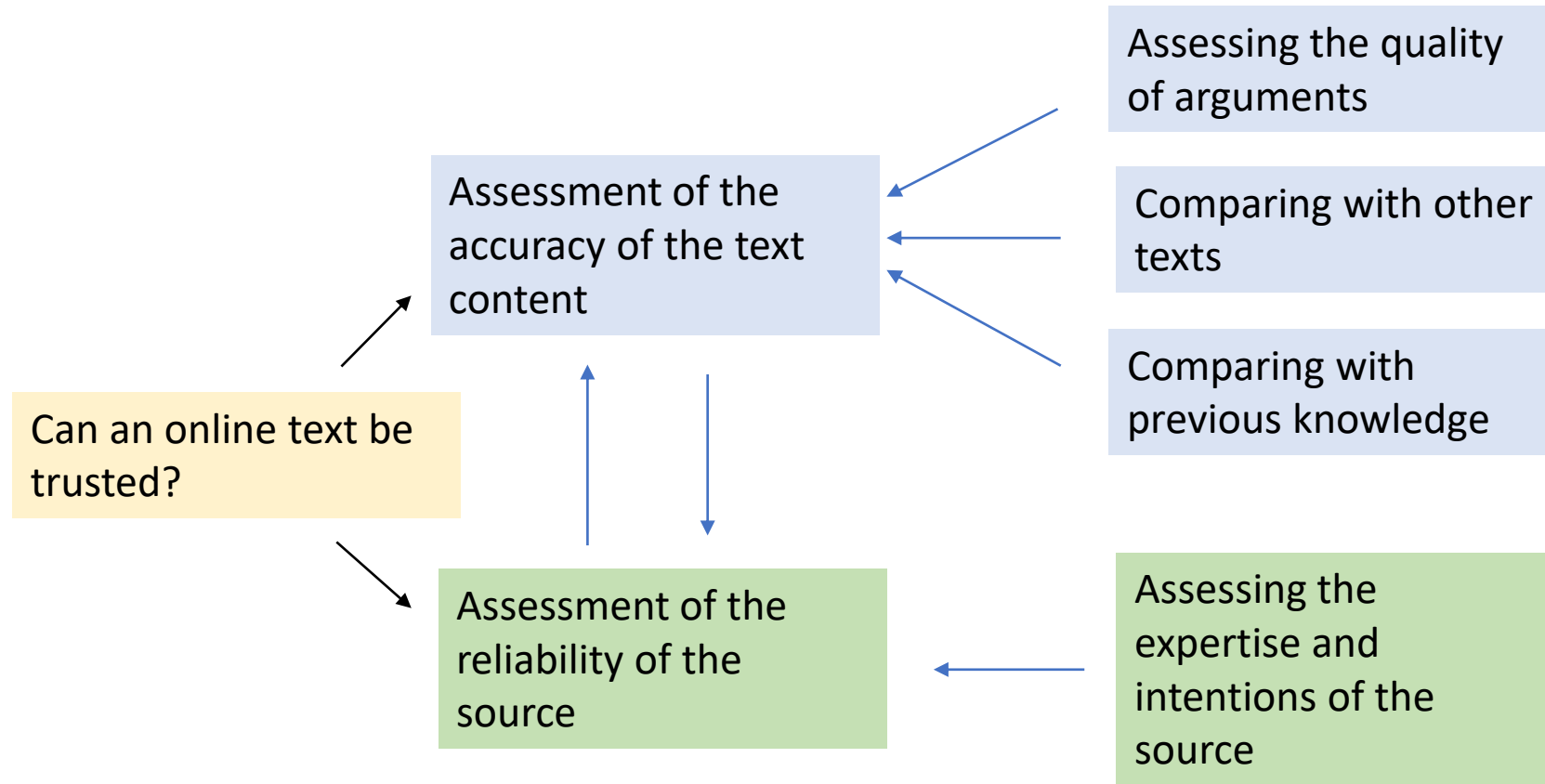
3 key questions



When confronted with a claim in the online environment, ask three key questions:

1. Who is spreading the claim? Source?
2. What evidence has been presented to support the claim?
3. What do the other sources have to say about it?

Checking the trustworthiness of online texts



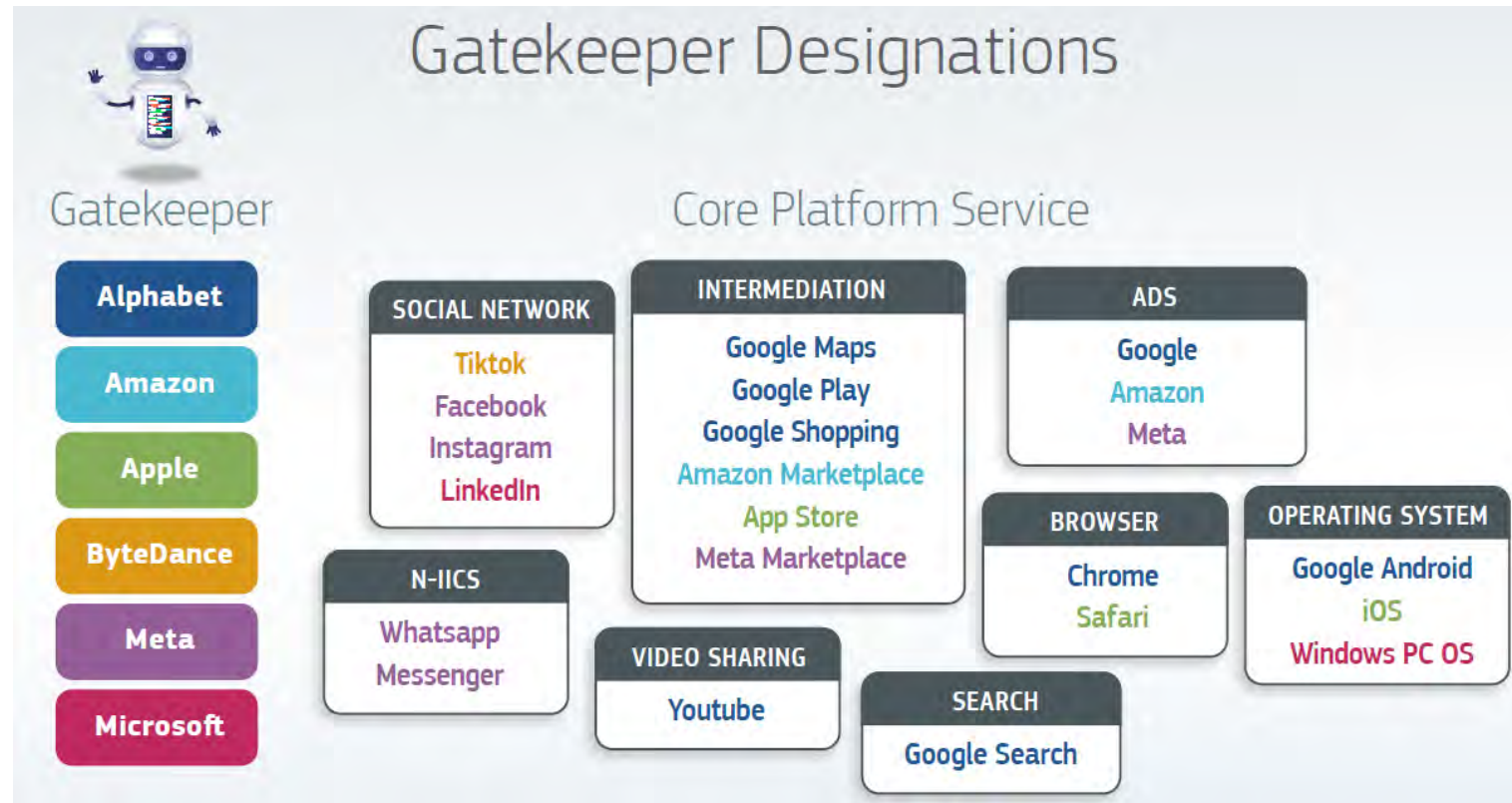
DSA – EU's Digital Services Act

- ❑ The Digital Markets Act's primary purpose is to ensure that digital platforms such as search engines, social networks, advertising services, and mobile app stores are **fair and contestable for both end users and businesses**.
- ❑ The DSA requires platforms to put in place measures to counter illegal goods, services or content online, such as a mechanism for users to flag such content and for platforms to cooperate with 'trusted flaggers'
- ❑ Impose conditions on reasonable advertising
- ❑ Enters into force on 17 February 2024



DMA – EU's Digital Market Act

- ❑ The DMA is a **competition law targeting the gatekeepers of the digital economy**. The Digital Services Act (DSA) obligates online platforms to be more transparent about how they collect data and outlines how to deal with illegal content and disinformation.
- ❑ The Digital Markets Act aims at preventing gatekeepers from imposing unfair conditions on businesses and end users and at ensuring the openness of important digital services.
- ❑ Enters in force March 2024



Generative Artificial Intelligence

- How to deal with GenAI?
- What are the pro's and con's for HR
- How could be use it in a positive way?
- The EU's first regulatory framework for AI, the AI Act, is currently in inter-institutional negotiation.



DigComp 2.2. - AI competence

- In order to be able to interact confidently, critically and safely with new and emerging technologies and AI, today's citizens would benefit from having a basic knowledge and understanding of AI and technologies.
- Awareness of the use of AI will help to better understand the challenges related to privacy, ethics and the rights of children and young people, for example



How to create a CV in 5 seconds?



Photo Boris Eldagsen - "Pseudomnesia/The Electrician" (Pseudomnesia/El electricista)

Won the creative open category at Sony World Photography Award (April 2023)

<https://www.bbc.com/news/entertainment-arts-65296763>

Generative Artificial Intelligence

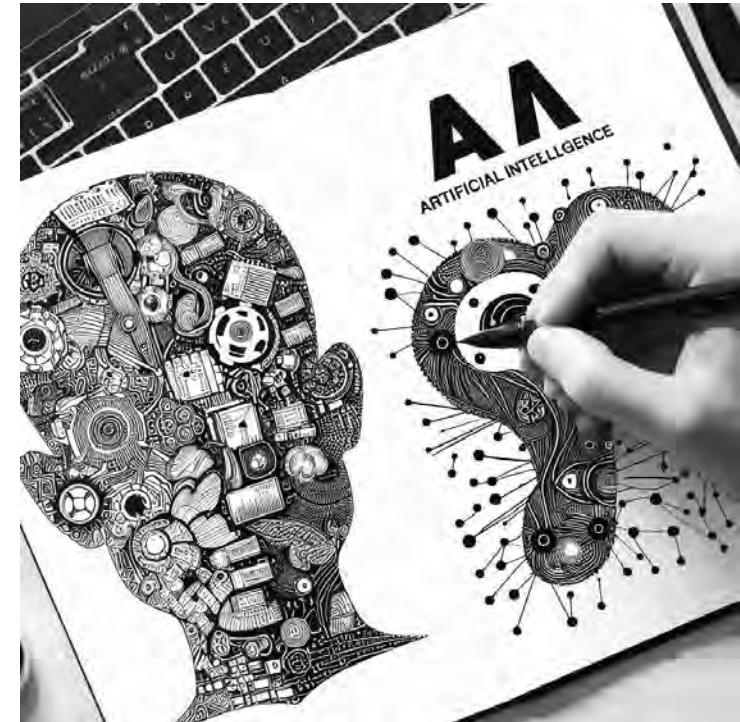
- ❑ A new generation of online available AI tools based on Generative AI models have the potential of severely impacting our work and our work methods.
 - ❑ They are computer programs which are designed to create new content that resembles human-made input.
 - ❑ **Generative AI refers to an AI application that generates content (e.g. text, images, code, and videos) in response to requests, prompts or questions written in natural language. No coding is needed!**
- ❑ ChatGPT
 - ❑ Dall-E
 - ❑ Midjourney
 - ❑ Bard AI
 - ❑ LaMDA
 - ❑ Perplexity
 - ❑ DeepL
 - ❑ Bing chat
 - ❑ Etc.



International Draft Guiding Principles for Organizations Developing Advanced AI systems [Commission](https://digital-strategy.ec.europa.eu/en/library/international-draft-guiding-principles-organizations-developing-advanced-ai-systems)
<https://digital-strategy.ec.europa.eu/en/library/international-draft-guiding-principles-organizations-developing-advanced-ai-systems>

Generative Artificial Intelligence & copyrights

- ❑ Generative AI models are trained from very large amounts of text, code, images, etc.
- ❑ They are often extracted from the Internet and usually without the permission of the owner.
- ❑ In fact, we do not know exactly what material was used for training.
- ❑ Several lawsuits have already been filed against GenAI companies challenging the use of copyrighted works as AI training material.



Generative Artificial Intelligence: INPUT issues

- ❑ Issues related to training in generative AI.
 - ❑ GenAI companies train their systems on datasets, which may include content collected by AI bots from the web.
 - ❑ This raises a number of copyright issues, such as the transparency and documentation of the content used by GenAIs to train AI.
- Solutions **to prevent scraping of copyrighted content** from the web. With regard to online content, several technical and legal approaches are under development or under consideration.
 - The **possibility of legally licensing content** for GenAI training materials, for example through collective rights management schemes. Some Member States have announced regulatory initiatives in this area.
 - **Developing case law:** rightsholders have brought a number of actions against GenAI companies that have already used their content as training material. Could such use be included in the copyright exceptions?
 - The international dimension of the territoriality of copyright.

Generative Artificial Intelligence: OUTPUT issues

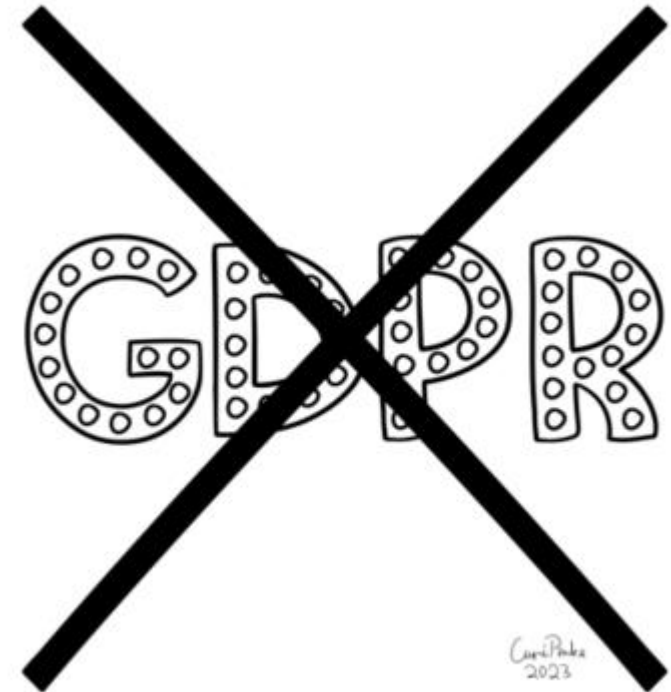
- The outputs of GenAI tools are generated from simple prompts or can be used to assist people in their innovative, creative, branding and other information and knowledge activities.
 - **Only human creations can be protected by copyrights at present!**
- **Copyright protection of GenAI outputs:** while some IP offices have issued guidelines on the criteria for copyright protection of AI-generated outputs, some are also considering ways to use other intellectual property rights to protect such works.
 - **Copyright protection against GenAI outputs:** some rights holders are also considering how to protect themselves from competition from outputs based on or imitating their works. In line with the transparency requirements considered in the context of the AI Act, this will require disclosure that the content was produced using AI, as well as detailed documentation of the copyrighted content used to train the GenAI tool.
 - **The use of GenAI tools for copyright infringing activities** and offences and their prevention, control and possible sanctions.

Generative Artificial Intelligence: Data protection

ChatGPT and many other apps are **not** compliant with the EU General Data Protection Regulation (GDPR), at least at the time of writing.

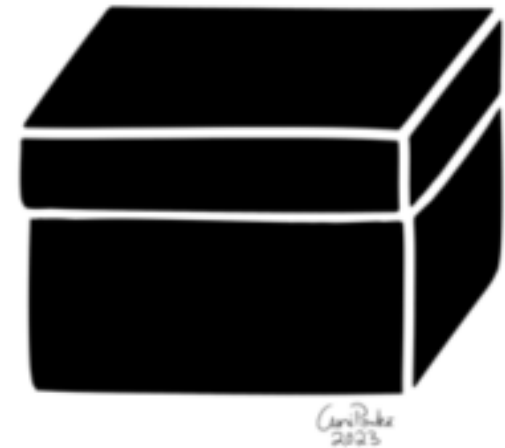
It is therefore important to remember that **you cannot share any confidential material**, such as client data, on these platforms

GPT systems may also be in conflict with other parts of the GDPR, based on which people have, for example, the right to be forgotten. However, deleting data from a GPT model is, at least for the time being, impossible.



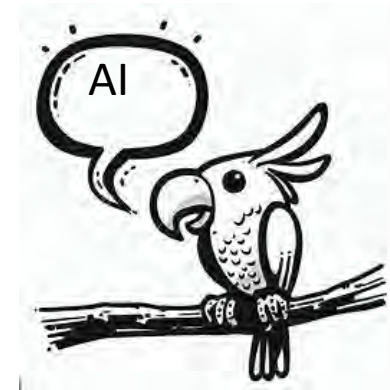
Generative Artificial Intelligence: Black box

- Artificial neural networks used by Generative Artificial Intelligence (GenAI) are generally "black boxes", i.e. their internal functioning cannot be studied.
- As a result, they are not "transparent" and it is not possible to determine how their outputs are generated. AI applications can therefore produce unexpected results.
- This opacity is also a key reason for the trust problems associated with GenAI models.
- For schools and business, the situation is difficult. Few have the necessary know-how or resources to check the potential risks of new software.



Generative Artificial Intelligence: Lack of understanding the real world

- An AI response can often seem so natural that it seems to understand the content of the text it produces.
- However, the GenAI does not understand the text it produces and often produces incoherent sentences.
- Therefore, all material produced by an AI must be viewed critically.
- AI systems produce text based on computational probabilities and have no understanding of the generated text.
- This is why GenAI applications are sometimes called "stochastic parrots" as they reproduce linguistic patterns they find in their training data (usually text extracted from the Internet) just as a parrot can imitate sounds without actually understanding what it is saying.



GenAI companies admit that GPTs invent things that don't exist. For example, the subsection of the ChatGPT public interface states that **"ChatGPT may produce inaccurate information about people, places or facts"**.

Generative Artificial Intelligence: Spreading mis- and disinformation

- GenAI tools have been trained on material from the internet, which often contains disinformation and discriminatory or otherwise questionable content.
- GenAI developers have made some efforts to weed these out, but in the absence of effective control mechanisms and regulations, biased or distorted material produced by GenAI is being disseminated.
- GenAIs may therefore produce and disseminate offensive and unethical material.
- While AI can produce outputs that appear remarkably insightful and useful, it can also invent "facts" that sound entirely plausible and incorporate them into its outputs. This phenomenon is often called **hallucination**.



Hallucination

AI & high quality fakes - e.g. deepfake videos



<https://www.buzzfeed.com/craigsilverman/obama-jordan-peelee-deepfake-video-debunk-buzzfeed>

Lesson learned in practise

An often-repeated concept in the debate on responsible AI development is
"human in the loop"

In practice, this means that ultimately only humans can both control the system and approve its outputs and use.



Generative Artificial Intelligence benefits for HR (Bing chat)

Here are some possible points for your presentation on the added value of generative AI for human resources. GenAi

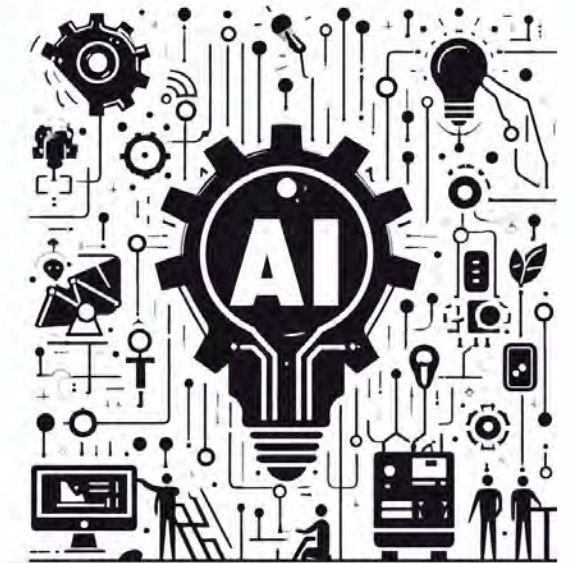
- **can create new content** or data based on existing data, such as text, images, audio, video, or code.
- **can save time and resources** for HR professionals by automating tasks that are repetitive, tedious, or require a lot of data processing. For example, GenAi can help draft job descriptions, create training materials, generate feedback, or summarize employee data.
- **can improve the accuracy and quality of HR outputs** by reducing human errors and biases. For example, generative AI can help create job descriptions that are more inclusive, relevant, and consistent.



Generative Artificial Intelligence benefits for HR 2

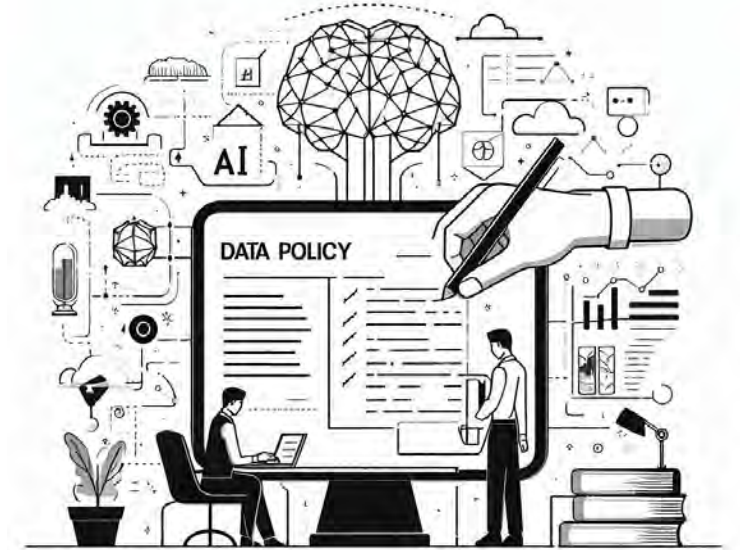
GenAi

- **can enhance the creativity and innovation** of HR professionals by providing new ideas, insights, and solutions (E.g. generate questions, scenarios, or exercises for interviews, assessments, or training.)
- **can foster closer human connections** and collaboration among HR professionals, managers, and employees by facilitating communication, feedback, and learning (e.g. create chatbots that can answer queries, provide guidance, or offer support.).
- **can empower HR professionals**, managers, and employees to leverage their own data and capabilities to improve their performance and development. (e.g. help to provide data-driven insights and recommendations that can help optimize HR strategies, processes, and outcomes)



Assess the AI capacity of your organisation

- ❑ Identify the skills, **roles and job descriptions** required by AI systems.
- ❑ Establish a **data policy** that defines how data will be collected and processed in AI systems.
- ❑ Establish criteria for AI systems in your organisation's **procurement policy**. Set concrete requirements, for example on the auditability of algorithms.
- ❑ Objectively assess **the appropriateness** of different technology options.
- ❑ Consider how you will set **objectives** for the AI system. Who will define the objectives and in what process? **Will staff be able to influence them?** – **human agency!**



Instructions to staff

- ❑ Staff must **never share any information** that is not already in the public domain, nor personal data, with an online available generative AI application.
- ❑ Staff should always **critically assess** any response produced by an online available generative AI model for potential biases and factually inaccurate information
- ❑ Staff should **never directly copy** the output of a GenAi model in public documents / emails **without verification**
- ❑ Staff should **never rely** on online available GenAi models of critical and time-sensitive processes.



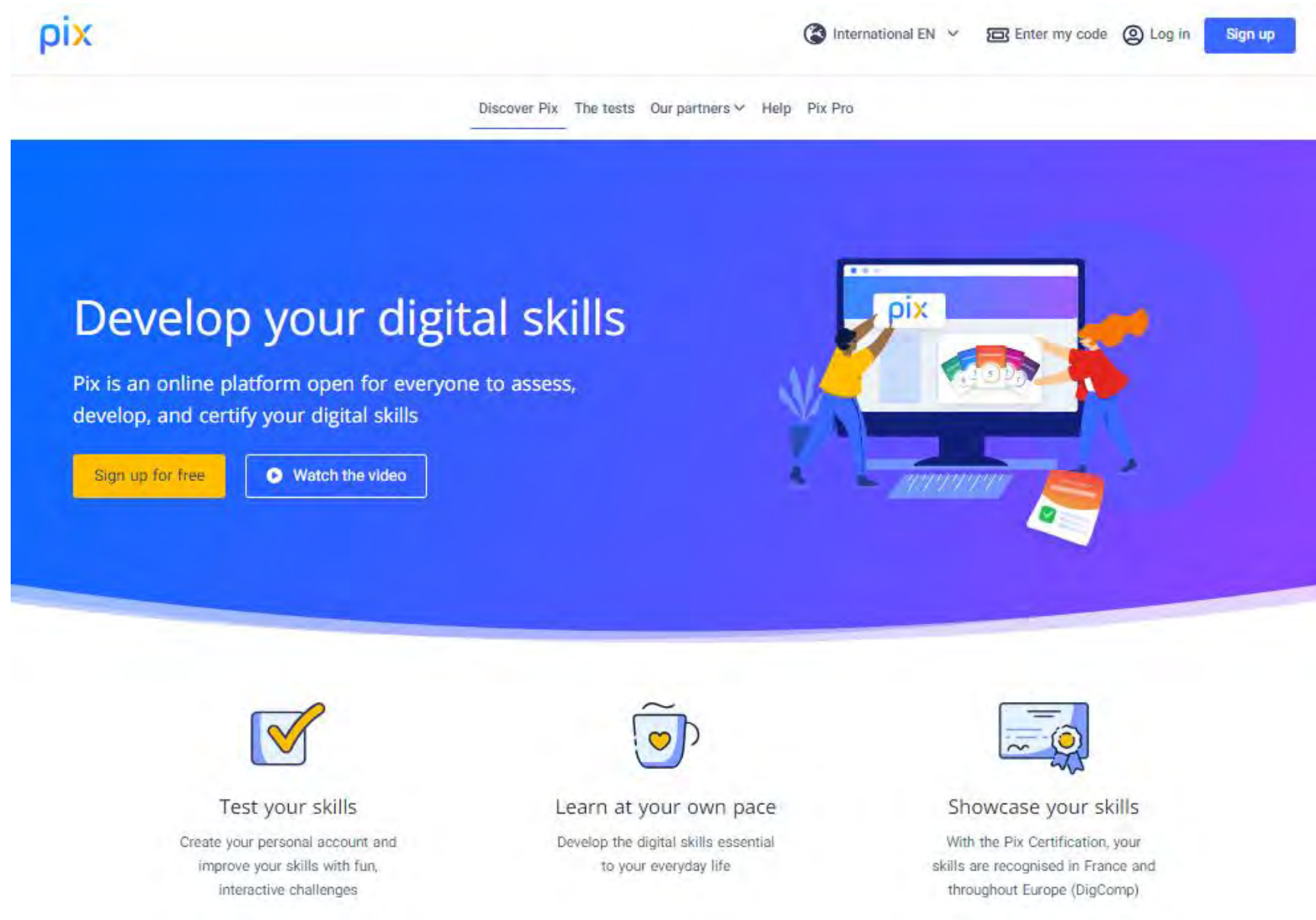
Which steps to take to act responsibly and effectively in digital landscape ?



Assessment tools

- French Pix-platform is assessing the real digital skills of person based on DigComp 2.2.

<https://pix.org/en>



The screenshot shows the Pix website homepage. At the top left is the 'pix' logo. On the top right, there are links for 'International EN', 'Enter my code', 'Log in', and a 'Sign up' button. Below the navigation bar, there are links for 'Discover Pix', 'The tests', 'Our partners', 'Help', and 'Pix Pro'. The main content area has a blue background with the heading 'Develop your digital skills'. Below this, it says 'Pix is an online platform open for everyone to assess, develop, and certify your digital skills'. There are two buttons: 'Sign up for free' and 'Watch the video'. To the right is an illustration of two people interacting with a large computer monitor displaying the Pix logo and a colorful bar chart. Below the main content area, there are three columns of information, each with an icon and a title:

- Test your skills**: Create your personal account and improve your skills with fun, interactive challenges. (Icon: checkmark in a box)
- Learn at your own pace**: Develop the digital skills essential to your everyday life. (Icon: coffee cup with a heart)
- Showcase your skills**: With the Pix Certification, your skills are recognised in France and throughout Europe (DigComp). (Icon: certificate with a ribbon)

Be careful of on-line scams and frauds

- Online crimes are exploding
 - Rapidly growing market
 - Low investments
 - Very low chance to get caught
- Several countries have
- centralised the reporting of the
- online scams



- Check the unknown site e.g. with Scamadviser before making any purchase!
- <https://www.scamadviser.com/>

References

- Debunking handbook: <https://www.climatechangecommunication.org/debunking-handbook-2020/>
- DigComp 2.2. <https://publications.jrc.ec.europa.eu/repository/handle/JRC128415>
- Digital Information Literacy Guide - <https://faktabaari.fi/dil/>
- Digiprofile test - <https://digiprofiletest.sitra.fi/>
- Consilium DSA information <https://www.consilium.europa.eu/en/infographics/digital-services-act/>
- Eurobarometer <https://europa.eu/eurobarometer/surveys/detail/2553>
- Faktabaari materials, www.faktabaari.fi
- Key Competences for Lifelong learning (2018). [https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32018H0604\(01\)&from=FI#d1e32-7-1](https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32018H0604(01)&from=FI#d1e32-7-1)
- Kivinen, K., Aslama, M., Havula, P., Härkönen, T., Kiili, C., Kivinen, E., Pönkä, H., Pörsti, J., Salo, M., Vahti, J., & Vuorikari, R. (2022). Digitaalinen informaatiolukutaito, Faktabaari, Helsinki , www.faktabaari.fi/dil
- Osborne, J., Pimentel, D., Alberts, B., Allchin, D., Barzilai, S., Bergstrom, C., Coffey, J., Donovan, B., Kivinen, K., Kozyreva, A., & Wineburg, S. (2022). Science Education in an Age of Misinformation. Stanford University, Stanford, CA. https://sciedandmisinfo.sites.stanford.edu/sites/g/files/sbiybj25316/files/media/file/science_education_in_an_age_of_misinformation.pdf
- Pix.org <https://pix.org/en>
- Scamadviser <https://www.scamadviser.com/>
- Vuorikari, R., Kluzer, S. and Punie, Y.(2022), DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg, *forthcoming*, ISBN 978-92-76-48882-8, doi:10.2760/115376, JRC128415. <https://publications.jrc.ec.europa.eu/repository/handle/JRC128415>

Thanks

kari@kivinen.net
Faktabaari.fi