

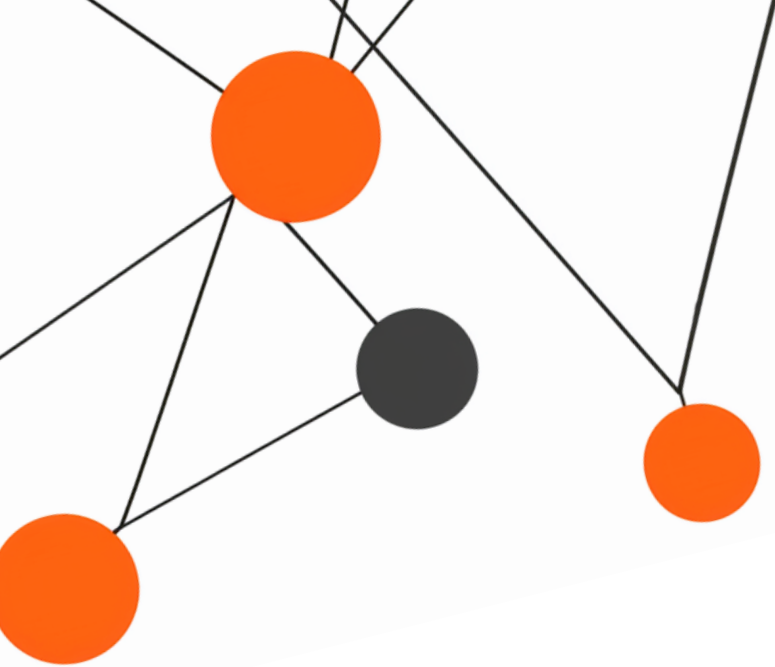
DIGITAL INFORMATION
LITERACY

AI GUIDE for Teachers



FaktaBaari **EDU**

MINNA ASLAMA HOROWITZ,
ARI HAASIO, PIPSA HAVULA,
KARI KIVINEN, JARI LARU,
MIIKA MIININ, HARTO PÖNKÄ,
PATI RUIZ, AKI SAARIAHO,
PASI SILANDER AND
TARMO TOIKKANEN.



Ai Guide **for Teachers**

ISBN 9978-952-65508-3-1 (softcover)

ISBN 978-952-65508-4-8 (PDF)

2025 Avoin yhteiskunta ry / Faktabaari EDU

Editor: Kari Kivinen

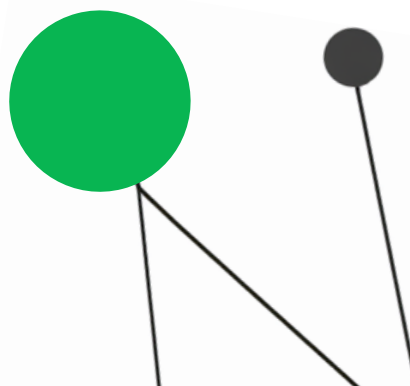
Authors: Minna Aslama Horowitz, Ari Haasio,
Pipsa Havula, Kari Kivinen, Jari Laru, Miika Miinin,
Tiina Härkönen, Harto Pönkä, Pati Ruiz, Aki Saariaho,
Pasi Silander and Tarmo Toikkanen.

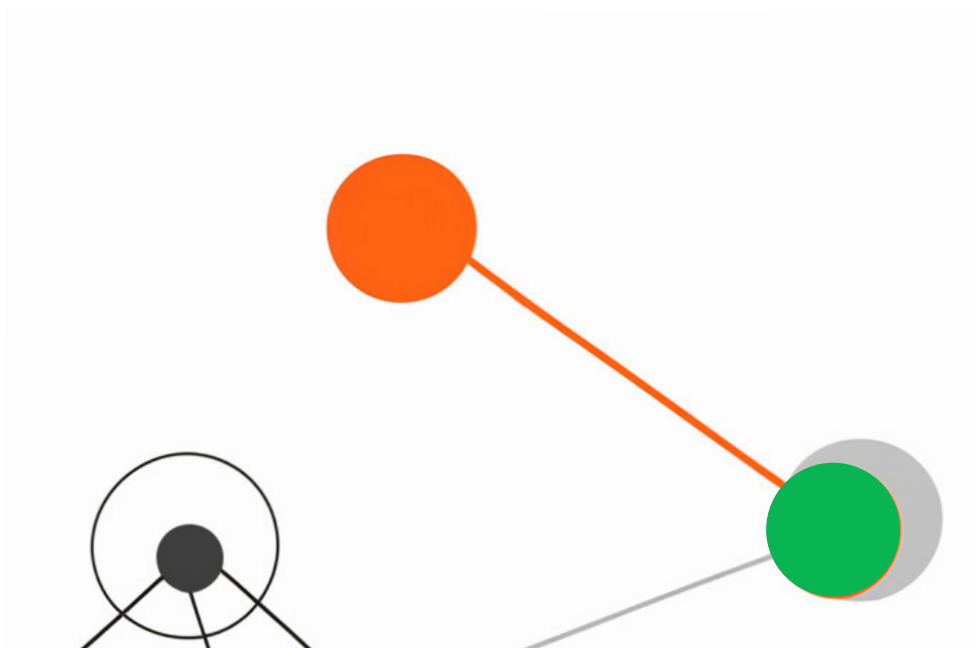
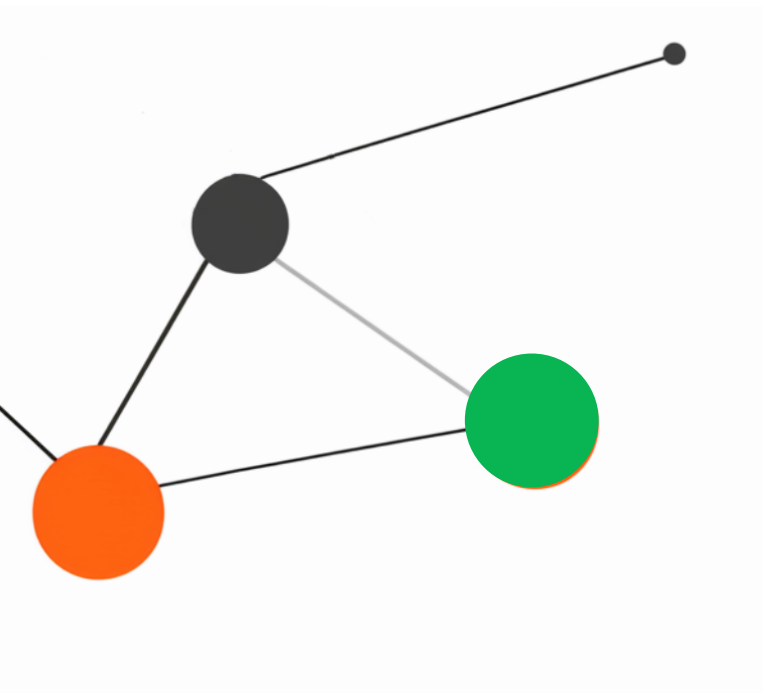
Quotes: Kivinen, K., Aslama Horowitz, M., Haasio,
A., Havula, P., Laru, J., Miinin, M., Pönkä, H., Ruiz,
P., Saariaho, A., Silander, P. & Toikkanen, T. (2025)
Tekoälyopas opettajille. Faktabaari, Helsinki

Layout and cover: Elina Salminen

Contents

1. Introduction	5
2. Getting Started with GenAI Tools	8
3. Artificial Intelligence as an epistemic change in education	10
4. How does Generative AI work?	13
5. Digital literacy, information resilience & teachers	19
6. Ethical challenges for educators in using AI	25
7. AI competency frameworks for teachers	30
8. Introducing AIAS - A framework for ethical integration of generative AI in educational assessment	34
9. What should teachers know about EU legislation on AI?	36
10. Teaching AI skills through "Generation AI" applications	39
11. Finnish "Generation AI" project	44
12. AI is changing pedagogy in Lappeenranta	46
13. Denial vs. techno-optimism: a teacher's perspective on AI	49
14. Deep fakes mimic reality	52
15. AI-based social media recommendation algorithms	55
16. AI - threat or opportunity for democracy?	65
17. AI and creativity - risk or opportunity?	67
18. AI in education case report: Crafting responsible use policies for emerging technologies in US school districts	70
19. GenAI, copyright and teachers	73
20. Data protection in AI applications	75
21. Appendix: Faktabaari guidelines for students	81





1. Introduction

KARI KIVINEN, FAKTABAARI

THE AI GUIDE is a follow-up to the Digital Information Literacy - AI Guide¹ published by Faktabaari in 2024. The aim of the publication is to give education providers, teachers and learners basic information about the functioning of AI, its capabilities, limitations, ethical challenges, benefits, misuse, AI pedagogy and AI literacy.

The aim is to clarify the educational perspective:

- **WHAT** is AI and how does it work (AI as a subject for learning).
- **HOW** AI can be used in a responsible and ethical way to support teaching and promote student learning (AI as a tool for teachers).
- **HOW** is AI changing society (Social impact of AI on the labour market, knowledge management, environment, communication, future professions, etc.).
- **WHY** is there a need for a healthy critical approach to AI outputs (The challenges of AI).

The guide has been written as clearly as possible to make it easy to read and understand. References will allow anyone to find additional information.

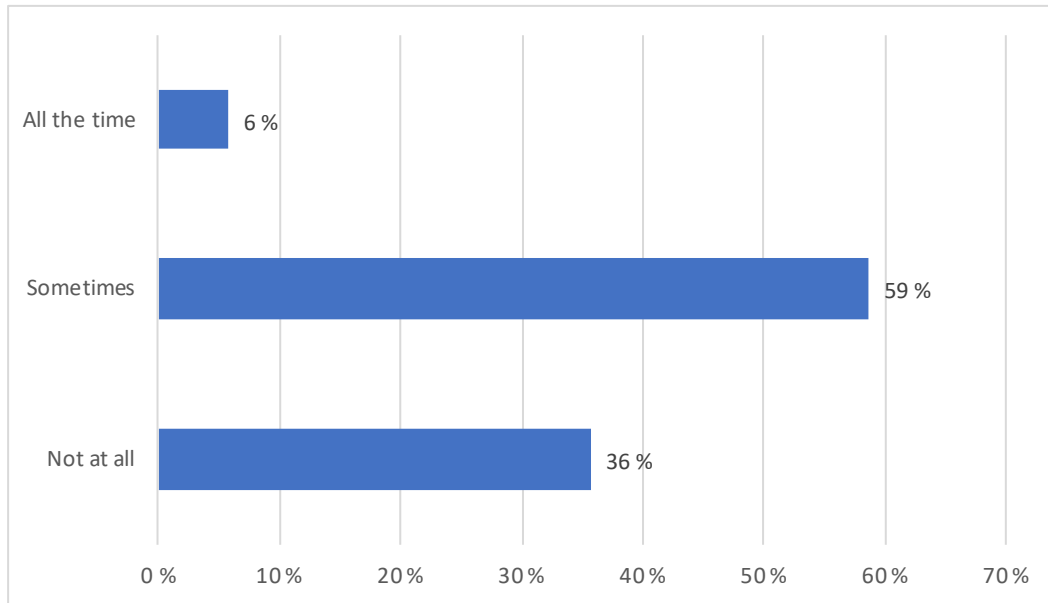
As AI technology is still in a state of flux and development, we have tried to select information that will stand the test of time.

The collection of articles is based on several international recommendations on AI in education published by the EU^{2 3 4}, UNESCO^{5 6 7}, OECD⁸, UNICEF⁹ and others.

In recent years, AI technology has advanced so rapidly that it has been difficult for education authorities to keep up with developments and provide updated guidance and in-service training for teachers. Teachers have therefore found themselves in a difficult position: How should they teach students to use AI technologies safely and responsibly when schools do not have clear guidelines on which software and services are safe to use and how to teach new skills within a partly outdated curriculum?

The relationship between education and AI is changing rapidly. AI developments and new easy-to-use AI systems - in particular generative AI - have already had a significant impact on teaching and learning. As machine intelligence increases and labour markets

Do you use AI tools at your work in the school? N= 331



and economies change, education systems will need to adapt to equip students with the knowledge and skills needed for the jobs of the future.

These developments require education and training providers to update their curricula and teachers to upgrade their skills to make use of new technologies.

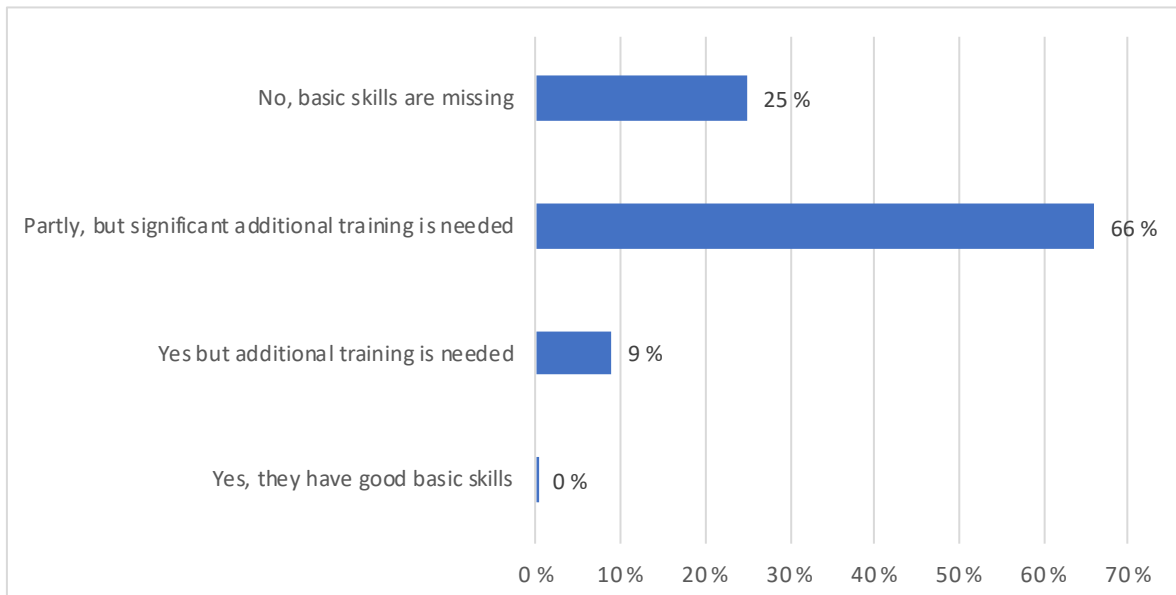
Use of AI tools by teachers

In autumn 2024 during an EUIPO-organised AI training programme, a poll about the AI tools

used among educators revealed a nuanced picture. The 331 teachers, from diverse European backgrounds, indicated that while AI is gradually being integrated into pedagogical practices, widespread adoption is still nascent. A significant proportion of respondents (36%) reported no use of AI tools in their teaching, while 59% indicated occasional usage. Notably, only 6% of teachers reported daily integration of AI tools into their educational activities.

The reason is probably that many teachers feel they do not have enough basic knowledge to implement AI.

Do teachers have sufficient AI skills? N=100, November 2024



In another AI training organised for Finnish teachers, none of the one hundred participants felt that teachers had sufficient AI skills, and as many as 75% of the teachers expressed the need for further training.

Human intelligence + artificial intelligence = augmented intelligence

Augmented intelligence refers to the use of technology that combines the strengths of human intelligence and machine intelligence. It does not replace human intelligence but rather extends and supports it.

The goal of AI is to help individuals and organisations make better decisions by harnessing the computing power and data analysis power of computers¹⁰. In education, too, the best results are achieved through collaboration between humans and AI.

AI vertigo

Sitra, an independent Finnish foundation dedicated to fostering sustainable well-being and driving innovation, has shared the term “AI vertigo” to describe¹¹.

Faktabaari’s AI guide offers an antidote for teachers suffering from AI dizziness. We have gathered background information on AI, summaries of international frameworks and recommendations, practical experiences of teachers who have used AI, and fresh reflections from experts from different perspectives.

1 Digitaalinen informaatiolukutaito- tekoälyopas (2024) Faktabaari. www.faktabaari.fi/dil/ai

2 Vuorikari, R., Kluzer, S. and Punie, Y., 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, 2022, ISBN 978-92-76-48883-5, <https://doi.org/10.2760/490274>

3 European Commission, Directorate-General for Education, Youth, Sport and Culture, Final report of the Commission expert group on artificial intelligence and data in education and training - A executive summary, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/65087>

4 European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

5 Miao, Fengschun, Cukurova & Mutlu (2024) AI competency framework for teachers. UNESCO <https://www.unesco.org/en/articles/ai-competency-framework-teachers>

6 UNESCO (2023) Guidance for generative AI in education and research. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

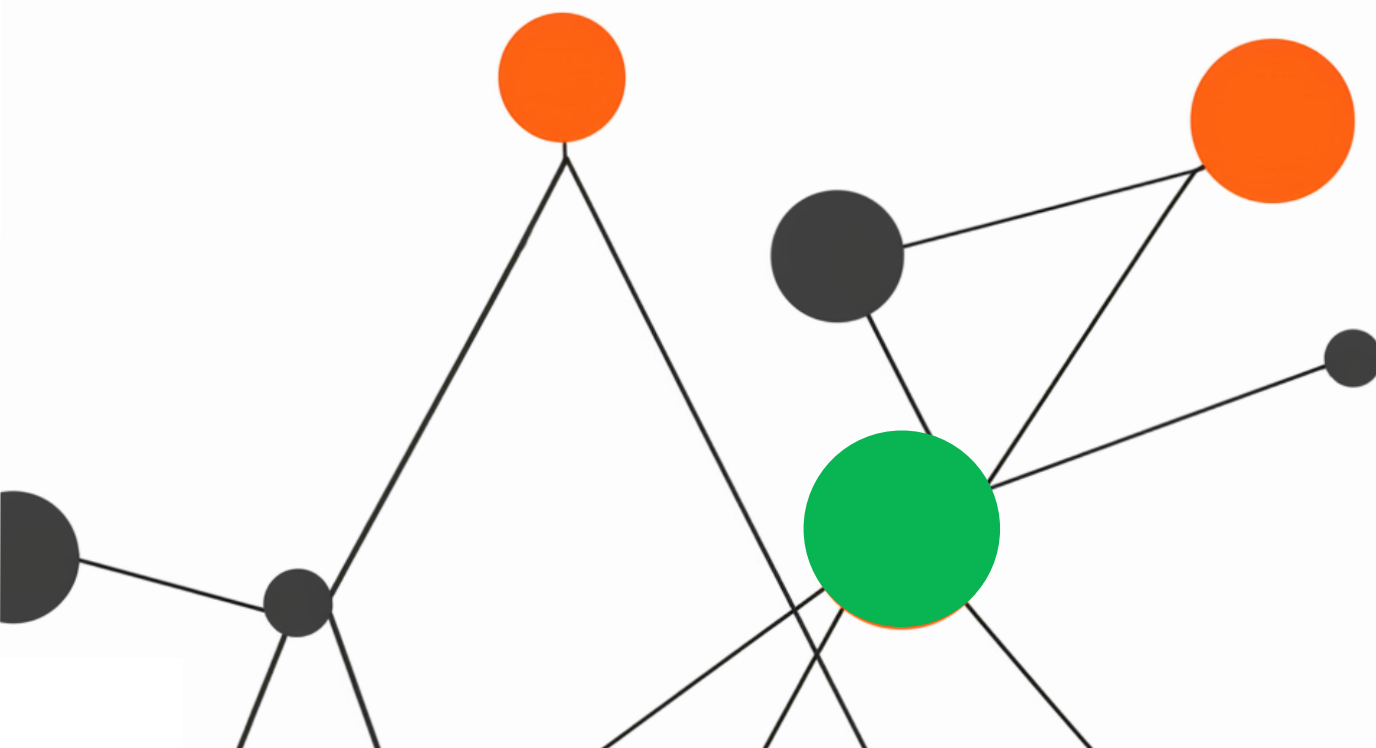
7 Fengchun, M. (2023). Guidance for generative AI in education and research. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>

8 OECD (2024) https://comite-etica.upc.edu/ca/actualitat/media/8-sti-explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_final_v3.pdf

9 UNICEF (uploaded 19.12.2023). Policy guidance on AI for children. <https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children>

10 Cf. Lari Numminen (uploaded 31.12.2024). Mitä on tukiäly? <https://www.finnishup.com/mika-on-tukiaily/>

11 Vuoden 2025 puheenaiheet (uploaded 31.12.2024) Sitra <https://www.sitra.fi/artikkelit/vuoden-2025-puheenaiheet-pida-naita-silmalla/>



2. Getting started with GenAI tools

USING GENERATIVE ARTIFICIAL intelligence (GenAI) technologies is easy and fun. They work directly in your browser.

First, choose the programme you want to use.

Here's a list of six AI services that are currently free:

GenAI	URL	Comment
OpenAi ChatGPT	https://chatgpt.com/	Best known, pioneer, versatile
DuckDuckGo	https://duckduckgo.com/aichat	DuckDuckGo AI Chat's free beta version allows you to conduct anonymous chats with third-party AI models.
Google Gemini	https://gemini.google.com	Fast, constantly evolving
Microsoft Copilot	https://copilot.cloud.microsoft	Also available as part of the Bing search engine. You can practise creating images e.g. with: https://www.bing.com/images/create
Perplexity	https://perplexity.ai	Can suggest sources, factual
Pi	https://pi.ai	Polite, user-friendly, good voice readers

Different programmes have their own strengths. Try out the different options and find the one that suits you most. All AI services also have paid versions that are more versatile and powerful. There are also a wide range of AI applications for music creation, sound cloning, podcasting, video editing, etc. which we will not cover here.

Open the programme and log in.

In most cases, you will need a free account.

Look for the input field (prompt) in the service:

This is where you write your question or task.

Enter the prompt:

Examples of text feeds:

- Write a tanka poem using the following words: winter, snow, ice, sun and spring.
- Explain why tokens are important for generative AI to work.
- Write an educational story about a fox, a hedgehog and a crow for a five-year-old child.

- **Draw up a lesson plan for a primary 5 lesson on checking the sources of generative AI output.**

You can ask the GenAI to refine and improve the output by continuing the input chain until you are satisfied with the result. The more precisely you describe what you want the AI programme to do, the more likely the output will meet your expectations. Most services understand Finnish and

produce grammatically correct text. However, it is worth remembering that the "mother tongue" of AI services is mainly English.

Always check the AI output!

It is important to always check whether AI is producing output that is appropriate and useful.

Sometimes the results are incorrect or fictitious. That's fine if it's a poem or a fairy tale, but you need to be careful when writing factual texts.

AI's own instructions

Let's ask Google Gemini what kind of advice it gives to the novice AI user:

Prompt:

Create a list of the five most important guidelines for a novice user of a generative AI service.

Of course! Here are the five most important tips for a novice user of generative AI:

1. Be specific and clear in your questions:

- **The more specific** you are with your question, the more accurate the answer will be.
- **Use keywords** that best describe the information you want.
- **Be as specific as possible:** instead of asking "Tell me about AI", ask, for example, "What are the biggest challenges facing AI today?".

2. Try different wordings

- **The same question** can be asked in many different ways.
- **Try synonyms** and different sentence structures.
- **If you don't get** the answer you want, try rephrasing the question.

3. Check the reliability of the answers

- **Don't trust everything as it is:** AI can sometimes produce incorrect information.
- **AI may sometimes give some information that is not correct.** Verify information from other sources, especially in important decision-making situations.
- **Critical thinking skills** are important when using AI.

4. Take advantage of AI's strengths:

- **AI excels at retrieving** and synthesizing information.
- **It can help you write,** translate texts and create different types of content.
- **Don't be afraid** to experiment with different functions and find the ways to use AI that work best for you.

5. Be aware of the limitations of artificial intelligence:

- **AI is not human:** It does not understand context as well as humans and can make mistakes.
- **AI can't understand the same context** as a human and can't understand the same language as a human, so don't rely blindly on the answers it provides.
- **Do not let AI to make decisions for you.**

Remember: AI is an evolving technology, and its uses are constantly changing. Be open to trying new things and learn more about the possibilities of AI.

3. Artificial intelligence as an epistemic change in education

PASI SILANDER

THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE has far-reaching effects on the skills needed in different areas of society and on the structures of working life, and thus also on schools and education. Throughout history, new technologies have shaped society and economic structures, such as the printing press, the steam engine, electricity or semiconductor technology. They enabled the transition from an agricultural society to an industrial society and later to an information society and digitalisation. While these technologies have

"Can we still tell what is real and what is created by artificial intelligence? Do we have the means to distinguish the truth from the untruth?"

had a major impact on production and logistics, as well as on communication and the economy, none of them has really challenged our traditional conception of knowledge and knowing. That said, the printing press, the internet and digitalisation have indeed significantly increased the amount of information and our access to it.

The development of artificial intelligence challenges us and our understanding of knowledge and knowing in a whole new way. Generative AI constructs and represents information in a very human-like way, even though it is based on Large Language Models (LLMs), which, in effect, is probabilistic computation. However, it is incredible how well language models produce information, both verbal and visual. The key question is: what kind of information or knowledge will learners need in society or workplaces in the future? How will teaching in schools have to change to enable learners to acquire the skills required in this age of artificial intelligence?

The conception of knowledge describes our understanding of the origin of knowledge, its nature, permanence, variability, and how knowledge is created or produced. We are accustomed to treating knowledge as a well-founded true belief, often modelled on reality and produced by humans. The prevailing constructivist view of knowledge emphasises the cumulative nature of knowledge, either as the product of an individual or as the result of collaborative knowledge construction. Socio-cultural artefacts such as language

	Surface learning	Deep learning	Future learning
Goal:	Recalling facts	Understanding	Creating new solutions
Outcome:	Capability to apply information only in a narrow context, if at all	Capability to apply knowledge in various situations	Capability to create new solutions for various new situations
Methods:	Information acquisition	Collaborative knowledge construction	Co-creation and co-innovation
Focus:	Facts	Knowledge	Thinking skills and strategies, as well as innovation practices

Learning – including its objectives and methods – needs to be viewed in a new light with the development of artificial intelligence².

or mathematics are central to the production of knowledge. Knowledge is thought to be the result of human cognitive processes, of thinking.

However, although AI handles, processes and produces information, it does not have human-like understanding or awareness. AI cannot create meaning or interpret information in a human way, although generative AI does an excellent job of mimicking human-like behaviour in its outputs. Nevertheless, the information produced by AI is always based on machine learning models, algorithms and underlying data. AI does not have the human conscience or capability for ethical or moral reasoning; nor the conscience and capability for emotion and empathy. The responsibility for knowing – for using and interpreting the information generated by AI – lies with the human, the AI user.

Pedagogical change

The change in the conception of knowledge is inevitably followed by a pedagogical change, which must consider the new need for knowledge and skills, for knowing. The constructivist

conception of knowledge is strongly reflected in the mainstream of school pedagogy; the constructivist conception of learning, which sees learning as the result of either the individual or the collaborative construction of knowledge by learners.

"AI is an innovation in itself, but above all, AI is a platform for new innovations."

However, in the age of AI, a new kind of *epistemic flexibility*¹ is needed, which can be defined as a flexible way of combining different

ways of knowing, thinking strategies and various types of knowledge (such as that produced by AI), always in a way that is appropriate to the situation. Epistemic flexibility is closely related to *metacognitive skills*, i.e. the ability to monitor, control and reflect one's own thinking.

In the near term, AI will not eliminate the need to learn e.g. reading, arithmetic or the core concepts of different subjects. However, it will change the kind of knowledge construction and thinking skills learners need. For example, learners will need *computational thinking*³ and data processing skills as prerequisites for understanding AI and how it works. These skills are necessary to understand automated decision-making and the use of machine learning in areas such as social media, hybrid influencing, internet search services or banking services. Computational thinking relies heavily on abstract reasoning, linguistic problem solving and mathematical reasoning.

Artificial intelligence in schools

A school's ability to respond to the changes in a digitalised society is determined, among other things, by the education system's conception

"Are we raising learners who are programmed by AI, or learners who programme AI?"

of technology (the concept and understanding of artificial intelligence), the conception of learning, teaching practices and up-to-date know-how. By developing these elements, schools and education can navigate the age of AI and ensure that learners have up-to-date skills relevant to society and working life.

In schools, AI cannot only be treated as a new technology, but as a force affecting democracy and freedom of expression in societies. It should be seen as a tool for automated censorship, hybrid influence, and information production. A key challenge for schools is to equip learners with critical thinking, problem-solving, source criticism and creativity skills related to the use of AI. These skills are a prerequisite for active participation and ensuring freedom of expression and democracy. In addition, the ethics of AI and the ethical use of technology must be a central focus in school curriculum.

"AI skills are often new to us and learning them is at its best a shared journey between teachers and learners."

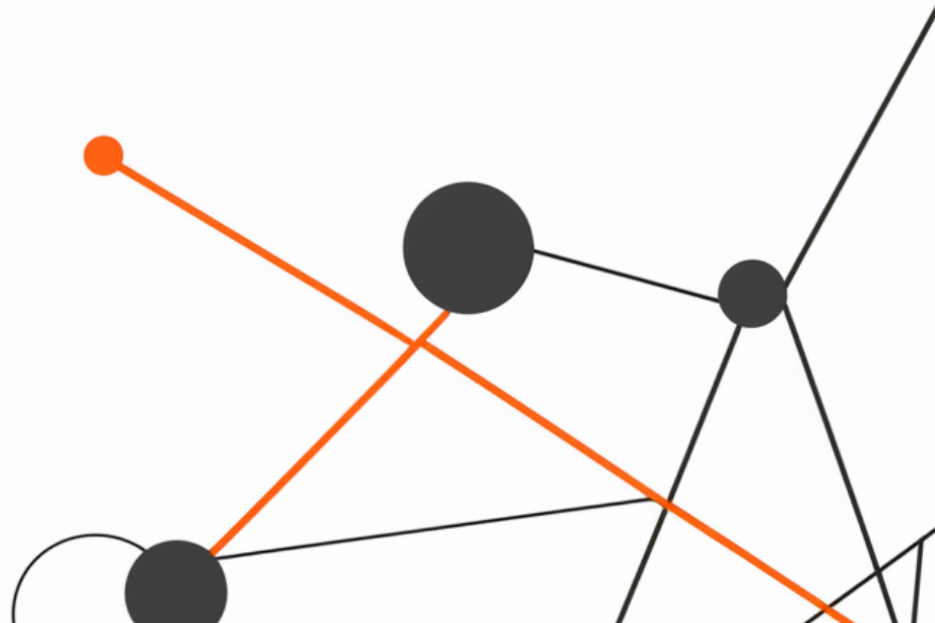
The use of AI in education requires that its users understand what AI is, what kinds of errors AI can make, and what types of biases AI outputs can have. It should also be noted that AI applications are not isolated from cultural and historical contexts, values and other contextual factors. Currently, AI applications are primarily based

on the commercial development work of large multinational corporations, the data they use, and the computing capacity of the cloud servers they provide. AI applications rarely consider the unique characteristics of small, less widely spoken languages or their cultural contexts.

¹ Markauskaite, L., & Goodyear, P. (2017). *Epistemic fluency: Innovation, knowledgeable action and actionable knowledge*. Springer Dordrecht.

² Silander, P., Riikonen, S., Seitamaa-Hakkarainen, P., & Hakkarainen, K. (2022). Learning Computational Thinking in Phenomena-Based Co-creation Projects: Perspectives from Finland. In *Computational thinking education in K-12: Artificial intelligence literacy and physical computing* (pp. 103-119). MIT press.

³ Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 366(1881), 3717.



4. How Generative AI works

KARI KIVINEN, FAKTABAARI

GENERATIVE AI (GENAI) uses Large Language Models (LLMs) and neural networks. It is a branch of AI that creates new content rather than simply analysing or classifying existing content or data.

The following definition of generative AI is derived from the Finnish National Agency for Education's version of the draft statement¹:

Generative AI refers to AI applications that can generate new content, such as text, images and videos. They are often based on machine learning and large-scale training data. Machine learning applications can learn new content from data. One form of generative AI is Large Language Models (LLM), which are used to generate and process human-like language.

While natural language is a common way to interact with GenAI, it is not the only way to interact with GenAI; inputs can also be, for example, code or images. GenAI generates new content based on a model created from training materials. Large language models can be generated either from public content on the Internet or from more limited, closed data sets.

In addition to large language models, generative AI can also be based on other models and techniques optimised for specific purposes, such as creating new music, videos or images. For example, so-called diffusion models focus on the creation and manipulation of images and sound.

Large language models

Large language models are trained on huge amounts of data. This data can include books, articles, code and other forms of written communication, as well as images and graphics.

Continuous data collection is made possible by data collectors – bots that collect data from the web. Nowadays, website operators can also opt out of data collection. It should be noted that data collection can also lead to the accumulation of inaccurate, discriminatory and biased information in training materials.

Language models use the collected data to learn the relationships between words and phrases and the meanings of different types of text. This enables them to perform a variety of tasks, such as translating languages, producing creative content and answering questions. As AI technology develops, these models will become more

sophisticated and offer opportunities for a wide range of applications.

Chatbots and virtual assistants use language models to provide users with responses to their input. Content creators use these models to produce articles, blog posts and marketing materials. In education, for example, they help teachers to create teaching materials and exercises.

Language models are built using neural networks, particularly the so-called transformer architecture. It is a deep learning method that provides an efficient way of processing e.g. sentences, text fragments and contexts. To make this possible, the text needs to be transformed into a numerical format that allows complex computations. This is called tokenisation.

Data points and tokens

Tokenisation is the process of converting text or images into a numerical format that computers can process. It involves breaking text into smaller pieces called tokens, which can be words, phrases or even single characters. Each token is given a numerical representation (embedding) that reflects its meaning in context. This method allows mathematical operations between sequences of words (tokens), for example by comparing their similarity. Contextualisation, which considers the occurrence of a sequence of words in the surrounding text, is crucial for the model to understand the meaning of a sentence or paragraph.

By breaking down text into tokens, models can better understand the underlying structure and meaning of the language.

Tokenisation splits the text into parts. Numerical representation (embedding) turns words into numbers (vectors). Contextualisation changes these vectors to reflect the meaning of the word in the context in which it occurs. This allows the computer to “understand” that the same word can mean different things in

different sentences. The vectors can also be used to calculate the similarity of words: vectors that are close together mean similar things.

A data point represents a single observation or event in a data set. A token is a smaller, indivisible unit of data extracted from a data point. In the sentence “The cat sat on the mat”, each word (“the”, “cat”, “sat”, “on”, “the”, “mat”) would be considered a token. However, tokenisation can be more nuanced. For example, the word “running” could be broken down into subword tokens such as “run” and “ing”. This approach, known as subword tokenisation, helps models deal with rare words or words that are not in their vocabulary.

Data points and tokens are the basic data units used to train and operate AI systems.

For example, a data point can be thought of as a complete recipe that lists all the ingredients, while a token can be thought of as a single ingredient in that recipe.

AI models work by processing many data points, each containing multiple tokens. The model analyses the relationships between these tokens to learn patterns and make predictions.

Tokens, data points and models are used to produce output. For example, GenAI tools can be asked to answer a question in the form of a text prompt that serves as input to a statistical model. The statistical model converts this prompt into numbers. The model then performs computations to predict the output based on statistical associations.

The best-known generative AI service is called ChatGPT, where GPT stands for “Generative Pre-trained Transformer”, which refers to the way ChatGPT is trained to process and understand large amounts of text data. This pre-training enables ChatGPT to produce human-like text in response to a variety of questions and requests. The model is a “transformer” because it uses a powerful neural network architecture.

The Financial Times¹ has made an excellent visual story about how the transformer works.

Language models are built using neural networks

Generative AI exists because of the transformer

This is how it works

By Visual Storytelling Team and Madhumita Murgia in London SEPTEMBER 12 2023

What material has been used to train language models?

According to OpenAI², the models used by the current ChatGPT have been developed using three primary sources of information:

1. **Information that is publicly available on the Internet,**
2. **Information that we partner with third parties to access, and**
3. **Information that our users or human trainers and researchers provide or generate.**

Unfortunately, however, these massive data sets often lack clear information about what they contain and where they come from³. AI companies generally do not disclose what data they have used to train their models. One reason is that they

want to protect their own competitive advantage. Another reason is that because data sets are bundled, packaged and distributed in a complex and opaque way, they are unlikely to know where all the data comes from.

The data used to train AI models is also heavily Western. More than 90% of the data sets analysed by the researchers came from Europe and North America, and less than 4% from Africa⁴. The dominance of English in the training data is partly explained by the fact that the Internet is still more than 90% English-speaking, and there are still many places on the planet where Internet connectivity is very poor or non-existent. This bias in the training material also has an impact on AI's performance and should always be viewed with a healthy dose of scepticism.

Holmes and Tuomi⁵ stress that the role of humans in the functioning of AI systems should

The Finnish Generation AI research project⁶ has developed an AI machine learning application for schools, designed specifically for use by primary school children. The project aims to enable learners, teachers and student teachers to understand AI/machine learning safety issues, technologies and models through experimentation and invention. All development work will be based on high quality international research.

In connection with the development of the teaching machine, school experiments have been carried out in Joensuu and Oulu, involving several hundred 4th and 9th grade students. The experience gained has been used to further develop the application. The teaching materials⁷ developed for the app can be found on the project website. More about this project elsewhere in this guide.

not be forgotten. Humans collect or curate the training data (e.g. images or text) used by AI, and humans write the algorithms or programme code underlying AI technologies and decide what to do with it.

Pre-training and neural network

During pre-training, the algorithms process a huge number of data points and tokens based on a large data mass.

A key element of generative AI is a neural network that mimics the general principles of neuronal function in the animal brain. A neural network is composed of interconnected nodes arranged in layers.

- **Input layer – receives and processes the raw data (e.g. an image, text or numerical values).**
- **Hidden layers – perform complex computations on the input data, extracting features and identifying patterns.**

Output layer – produces the network's output, which can take various forms.

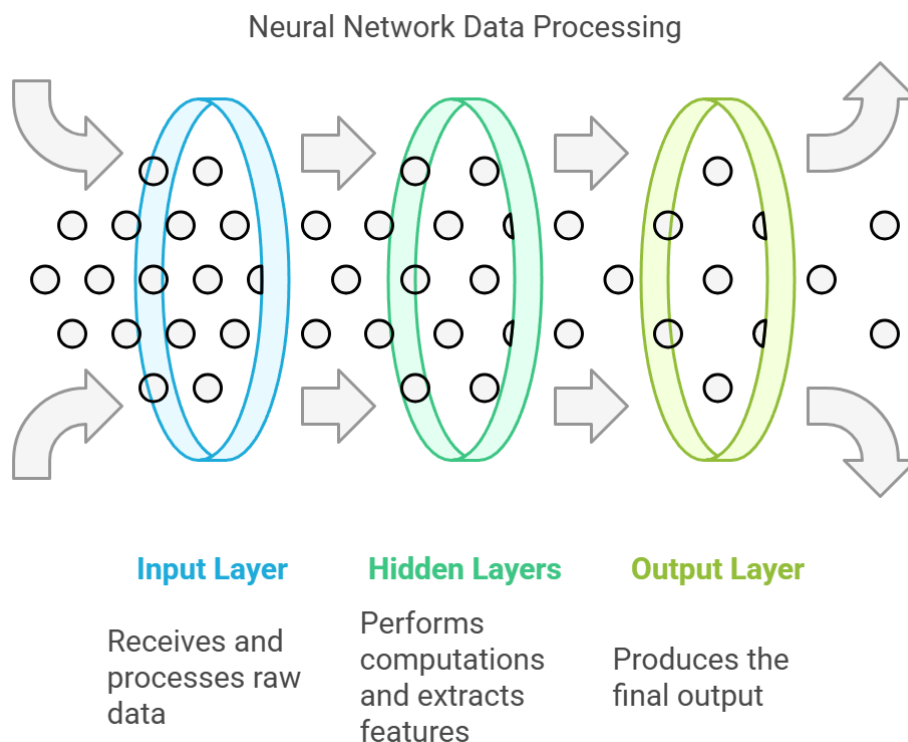
During training, the AI model learns the relationships and patterns within the data by

adjusting the strength of the connections between nodes, called weights. The training process involves continually adjusting these weights to minimise errors and improve the model's ability to predict or generate content. Optimisation algorithms are commonly used to adjust the weights, the most important of which is currently the "GTP", after which ChatGPT is named.

Once the AI model has been sufficiently trained, it is able to generate content based on the given input. Based on the patterns and relationships it learns from the training data, the model produces a consistent and contextually appropriate output for the training data that was used.

Generative AI utilises deep learning models, which are part of machine learning models. Machine learning refers to the ability of a system to learn autonomously from the data it is given, without a human determining all its functions. Deep learning allows AI models to learn to recognise patterns and connections in massive amounts of data, such as images or text. They can then generate new, similar data that replicates the

Neural networks are made up of interconnected nodes, arranged in layers.



Made with Napkin

features present in the original data. This differs from traditional rule-based AI systems, which follow predefined rules and are unable to learn or improve over time.

Training AI models requires huge amounts of data, high-performance technology and financial resources. Only very large companies can develop these models. Fortunately, these models can be used for many different purposes.

Although generative models are powerful, they also have limitations. For example, they can produce erroneous conclusions and biased content.

Fine-tuning

Fine-tuning aims to adapt a foundational model to a specific task. The model is trained on a smaller and limited set of data related to the desired application. Fine-tuning improves the performance of the model in the given task domain and makes the model more cost-effective to use.

Fine-tuning enables the development of GenAI applications for a wide range of purposes. Indeed, dozens of new AI applications are appearing on the market every week for a wide variety of purposes.

Input, prompt or feed

An Input is a command given by the user to the AI. It can be text, sound, image or programme code. Based on the input, the AI produces the desired result. The more precise and detailed the input, the better the result.

It is the inputs that influence the outcome produced by the generative AI. Skilful design of inputs is therefore an important skill to practice. Fortunately, there are many instructions and examples available on the Internet, and many AI services also provide tools and example feeds. You can also ask AI services directly for advice on how to create a better feed. In most cases, it is not a good idea to settle for the first result, but to work on it further by asking the AI service to refine and improve the output.

The difference between generative AI services and search engines

It is very important to recognise that “traditional” search engines and generative AI software serve different needs and provide different services to their users. Search engines search for existing information, while GenAI services create new content.

Search engines use algorithms that analyse page content, number of links and other factors to determine the relevance and order of pages in search results. They search and organise data on the Internet based on the search terms entered by the user.

Search results are based on both the user’s previous searches and the recommendations of the search engine’s algorithm.

In most cases, the information seeker will receive thousands or even millions of direct links to data on the Internet. It is the user’s responsibility to choose the one that best suits their information needs. On the positive side, the user can directly assess the reliability of the information source. Unfortunately, search engines such as Google do not use transparent search logic, and organise and censor results based on the user’s profile. Paid or sponsored results often come first in search results, as search engines are driven by commercial interests.

More recently, GenAI services have started to propose a combination of these two different approaches. For example, Google Gemini (version 1.5 Flash) offers the possibility to “check” GenAI results with the Google search engine. This function can be found in the three-dot menu below the answer (“Check the answer carefully”).

If you use a “normal” search engine to find information on, say, Faktabaari, you will be presented with millions of very different results (Bing: over 15 million hits), including criticisms of Faktabaari. Asking ChatGPT’s online search function “What is Faktabaari” will return a short

Training AI models requires huge amounts of data, high-performance technology and financial resources.

summary of Faktabaari. One may wonder why particular sentences are chosen and whether the ChatGPT summary gives a realistic and comprehensive picture of Faktabaari's activities.

The operating mechanism of generative AI applications creates new content from large amounts of data. The user usually receives only one answer at a time, based on computed probabilities according to the process described above. On the other hand, the user can receive additional information by submitting finetuning

questions and prompts and the results can be further processed and refined.

Although the answer produced by the GenAI service often has a high degree of confidence, it may be incorrect. GenAI models do not understand

the text they produce but are based on statistical models and training data. Therefore, their outputs may be incorrect or misleading. This highlights the need for source criticism when checking the outputs of GenAI services. So, for example, under

the Gemini input field it says "Gemini can make mistakes, even when it comes to information about people, so it's important to double-check its responses."

Useful links

How Chatbots and Large Language Models Work video <https://youtu.be/X-AWafSFCHQ>

Generative AI exists because of the transformer (uploaded 18.1.2025) FT. <https://ig.ft.com/generative-ai/>

1 Tekoäly varhaiskasvatuksessa ja koulutuksessa - lainsäädäntö ja suosituksset (uploaded 17.10.2024 lausuntopalvelu.fi) <https://www.lausuntopalvelu.fi/Proposal/Participation?proposald=a0d6af03-67e1-4ec7-9269-fab75bb05807>

2 OpenAI (Haettu 13.12.2024) <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed>

3 MIT Technology Review (Haettu 18.12.2024) <https://www.technologyreview.com/2024/12/18/1108796/this-is-where-the-data-to-build-ai-comes-from/>

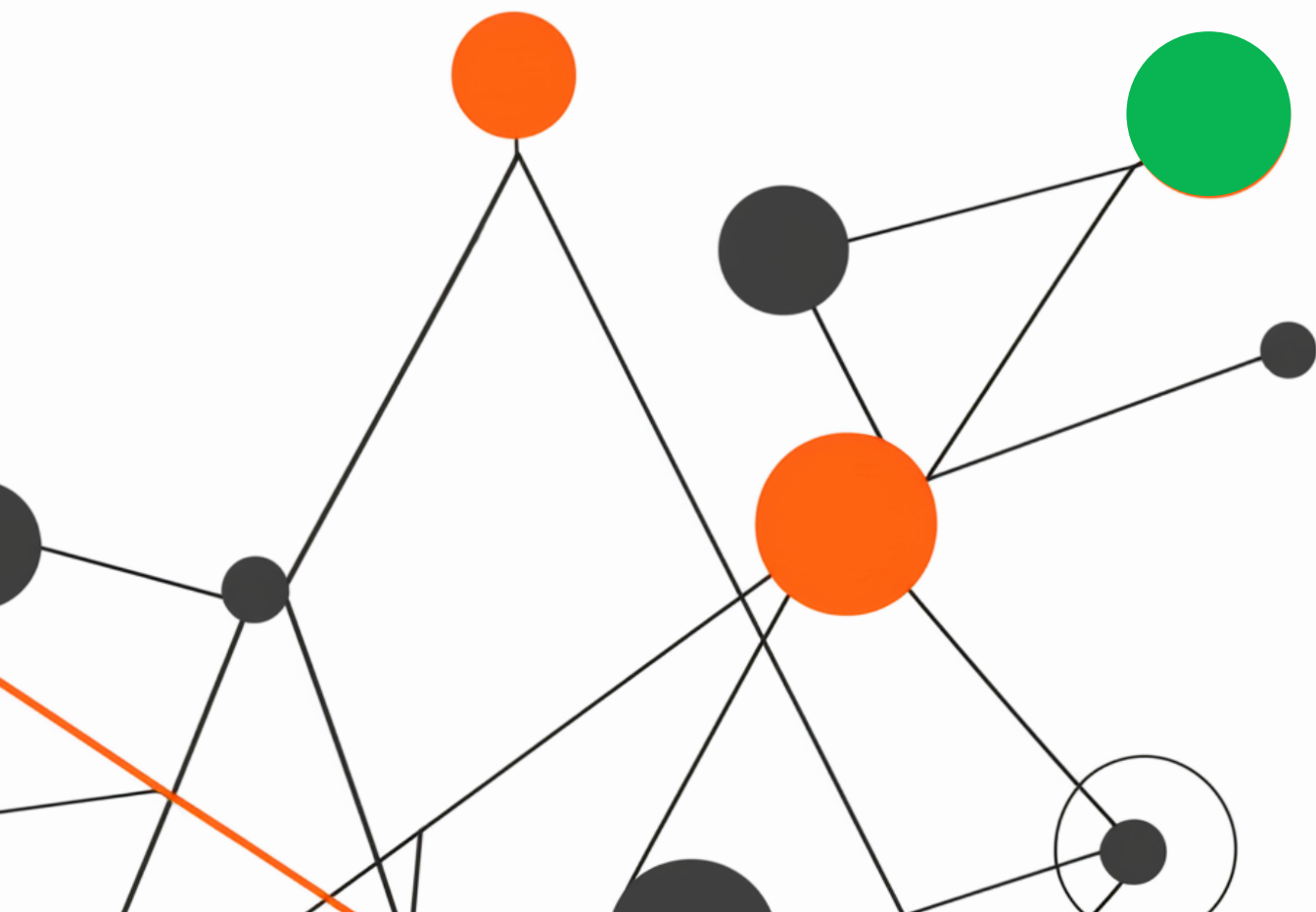
4 MIT Technology Review (Haettu 18.12.2024) <https://www.technologyreview.com/2024/12/18/1108796/this-is-where-the-data-to-build-ai-comes-from/>

5 Holmes, Wayne and Tuomi, Ilkka (2022) State of the art and practice in AI in education. Special Issue: Futures of artificial intelligence in education. European Journal of Education. Volume 57, Issue 4. <https://onlinelibrary.wiley.com/doi/10.1111/ejed.12533>

6 Generation AI tutkimushanke <https://www.generation-ai-stn.fi/>

7 Generation AI teaching machine materials <https://www.generation-ai-stn.fi/materiaalit>

Although the answer provided by the GenAI service is often very reliable, it may be incorrect.



5. Digital literacy, information resilience & teachers

KARI KIVINEN, FAKTABAARI

DIGITALISATION HAS BECOME part of our everyday lives, and its impact extends to education. So, what are these digital skills?

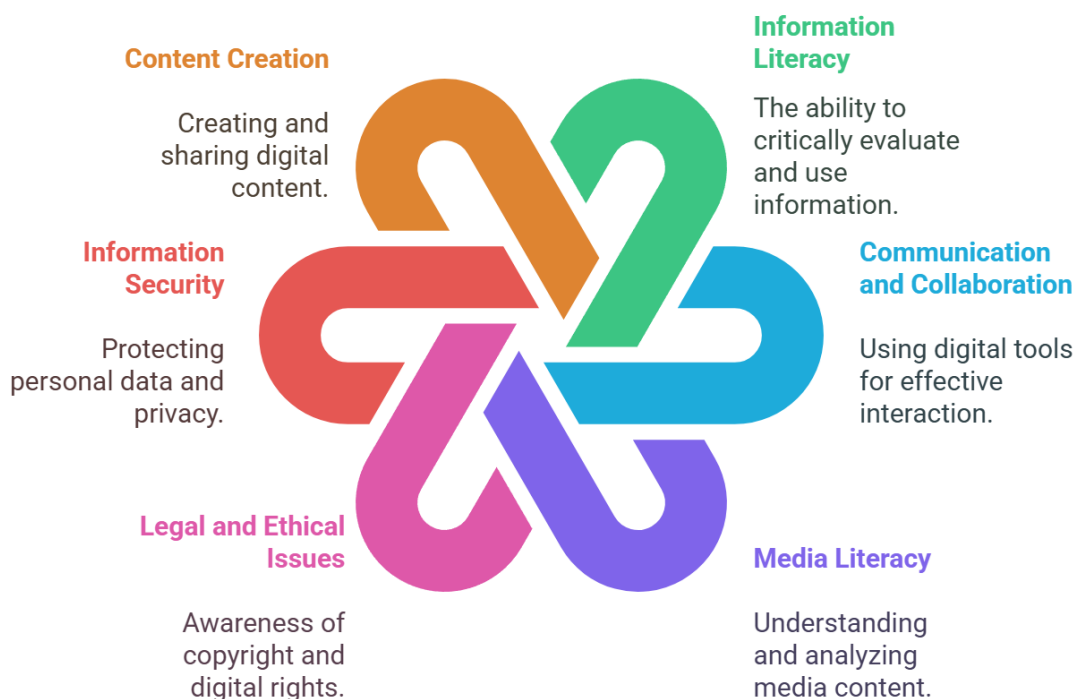
Digital literacy is not limited to the use of a computer or smartphone, but can include the following areas:

The European Digital Competence DigComp 2.2.1 Framework provides concrete examples of the knowledge, skills and attitudes citizens need to assess digital content:

Knowledge

- Be aware that online environments contain all types of information and content including misinformation and disinformation.
- Know the importance of identifying who is behind information found on the Internet (e.g. on social media) and verifying it by checking multiple sources.

Essential Digital Skills for Citizens



Made with Napkin

Skills

- Know how to differentiate sponsored content from other content online.
- Know 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.
- Know how to find the author or the source of the information, to verify whether it is credible.

Attitudes

- 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.

Citizens' resilience against information interference

Over the past two years, the debate has increasingly focused on information and electoral influence via the Internet and on strengthening citizens' resilience.

The 2024 edition of the World Economic Forum's Global Risk Report² concludes that disinformation is the most serious threat in the short term (2 years), and the fifth most serious threat in the longer term (10 years). According to the report, disinformation can radically disrupt electoral processes and cause civil unrest and

clashes. It also leads to growing distrust of the media and government sources. It also deepens polarised views even in established societies.

There are many cautionary examples of such behaviour, the most recent and serious being Romania, which on 6

December 2024 became the first EU country to cancel elections due to foreign interference

Disinformation can disrupt electoral processes and cause unrest and polarisation.

following reports of data manipulation on TikTok.³

Disinformation spreads almost unhindered on social media platforms, even when borders are closed. Finland is not immune to the threat of disinformation.

An editorial in Finnish newspaper Ilta-Sanomat (4.8.2024)⁴ sums up perfectly

what should be done: "In addition to reliable media, fighting disinformation requires, among other things, state preparedness for the threat – and good media literacy education. At the forefront of the fight are vigilant citizens who are able to recognise attempts to influence them and do not get provoked when provoked."

But resilience does not happen by itself, and knowledge and skills need to be diligently updated, especially as AI technology develops. More is needed, such as:

- **Regulating platforms and monitoring and communicating regulations (e.g. EU digital and AI regulations are a good starting point);**
- **Research;**
- **Fighting against disinformation and cooperating internationally (e.g. Faktabaari⁵, NORDIS⁶, EDMO⁷);**
- **Reliable journalism and media self-regulation (e.g. Council for mass media in Finland⁸);**
- **Digital and AI education (in schools, libraries, workplaces, military, adult education, etc.).**

Digital information literacy

Digital literacy is therefore not just about technical infrastructure, technology, applications, devices or network connections. Faktabaari – in partnership with teachers – has chosen the following key area of action to promote:

Digitalisation is not just about technical infrastructure, technology, applications, devices or network connections.

- **Young people’s ability to critically evaluate online content and generative AI outputs;**
- **Understanding of the impact of algorithms on search results and recommendations;**
- **An awareness that our data constantly being collected and used for informational purposes;**
- **The ability to manage our privacy and digital security in the digital environment.**

Faktabaari presented a definition of digital information literacy in the DIL Guide⁹. It has since been expanded to include AI literacy as part of digital literacy:

Digital information literacy is the ability to access, manage, understand, integrate, communicate, evaluate, create, and disseminate information safely and appropriately through digital technologies.

It includes information and media literacy, as well as artificial intelligence and data literacy, which helps understanding of the digital information world in a broader sense.

Digital information literacy involves a dimension of active and civic engagement with the digital world and promotes active citizenship.

The guide also presented the Finnish curriculum’s definitions of multiple literacies and international literacies, as well as various online literacy skills and strategies.

Over the past two years, the online reflection on the three questions below has proved to be very useful.

When faced with unfamiliar online content, it is always a good idea to find the answers to three simple key questions before spending time exploring the content in more detail:

1. **Who is behind the information? (source)**
2. **What is the basis for this claim? Evidence?**
3. **What do other sources say?**

Teaching techniques¹⁰ of online reflection have been shown to be an effective way to promote digital source criticism and lateral reading skills among citizens¹¹. Lateral reading refers to the process of checking the background of online information (reliability of the author, facts, statistics, sources, etc.) from different sites and sources before delving into the text at hand.

Faktabaari has published¹² educational videos and teaching materials to support teachers. For example: Recommendation algorithms; Check 3 questions; Identify error; Lie & gossip; Try reverse image search; Lateral reading; etc.

According to the recommendations of the Finnish Critical Research project¹³, every young person needs critical literacy skills, i.e. the ability to search for information in different media, assess the reliability of texts and compare the information they contain. According to the study, critical literacy and basic reading skills should be supported in parallel and reading challenges should be addressed as early as possible. They have also produced tools for teaching critical online literacy¹⁴.

The Finnish framework for digital competence

The Finnish Ministry of Education published the “Digital competence in early childhood education and care, pre-primary education and basic education” frame¹⁵ in 2022, just before the global roll-out of generative AI.

The framework descriptors clarify with concrete examples how digital competency can be built up from early education to the end of comprehensive basic education. The aim is for education providers to update their digital strategies and plans and their curricula in line with the proposed new digital framework.

Digital competence and ICT competence as a whole consists of four main areas:

4. **Practical skills and personal production**
5. **Security and responsibility**

6. Information management and inquiry-based and creative work

7. Interaction

Concrete and illustrative descriptions encourage teachers to use digital environments in a variety of teaching activities. Digital environment refers to all digital solutions, services, devices and tools used in education and early childhood education – and more recently also AI services. The competency descriptions give an indication of what digital skills should be taught at which level. New technologies enable learners of all ages to play not only passive social media roles, but also active roles as actors, producers and researchers.

The teaching of digital skills will ultimately fall to teachers at all levels and in all subjects. It is therefore essential that education and early childhood education providers are responsible for ensuring that the equipment and services available to teachers comply with the EU Data Protection Regulation and national legislation. In addition, teachers should receive systematic in-service training and clear guidance on how to work safely and responsibly in digital learning environments.

PISA, digitalisation and mobile phone bans

According to PISA 2022¹⁶, 15-year-olds' maths, science and literacy skills declined in most OECD countries, including Finland.

The reasons for the decline in results have been the subject of a wide range of speculation¹⁷. Possible reasons for the drop in results have been put forward in the public debate, including

New technologies also enable learners to play an active role as actors, producers and researchers.

digitalisation, pupils with an immigrant background, pandemics, phenomenological learning, inclusion and self-direction.

The PISA study provides a rather limited view on the link between the use of digital devices and learning outcomes. Pupils have

experienced the use of devices as somewhat disruptive to learning, but on the other hand, moderate use of digital resources seems to have a positive link to learning outcomes¹⁸.

The debate on the use of mobile phones by children and young people has been heated, and their use in lessons has been and will be further restricted.

However, restricting the use of mobile phones does not mean that schools should not give all children and young people the opportunity to learn important digital skills, responsible use of the Internet and critical online literacy.

The debate should also consider whether Finnish teachers have sufficient basic knowledge to use digital environments and AI applications effectively to support teaching, and how to organise and fund the necessary systematic digital teacher training.

How can you develop your digital skills?

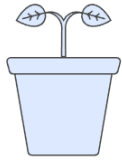
Learning digital skills, for example by testing and playing with different AI services, can open a new way of doing everyday tasks. For example, AI can be a fun way to make travel or holiday plans, effortlessly translate text from one language to another, and create great pictures and videos. They are powerful tools for planning lessons, selecting materials and brainstorming. Teachers can use their digital skills to inspire students to use their personal smart devices for more than just browsing and playing on social media.

Teachers have a key role to play in developing their students' digital skills. Digital skills are essential in the world of work and help students participate in social debate. New technologies can make learning more engaging and interactive.

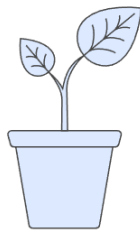
AI literacy

AI is a key emerging digital skill. It refers to an understanding of how AI works, its principles, concepts and applications, as well as an awareness of the limitations, implications and ethical considerations of using AI.

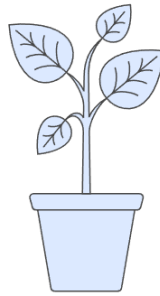
Developing student's digital skills



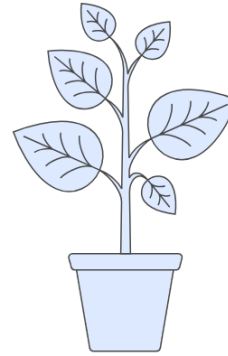
Try new tools



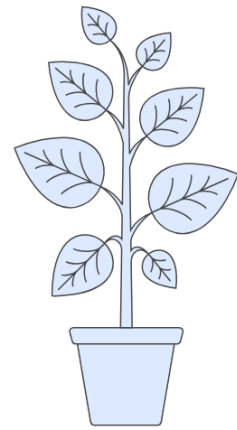
Teach critical thinking



Discuss security



Address ethical issues



Encourage creativity

Made with Napkin

Currently, the most recent and comprehensive definition of AI literacy is based on the definition presented by Bozkurt in autumn 2024¹⁹:

AI literacy is the comprehensive set of competencies, skills, and fluency required to understand, apply, and critically evaluate AI technologies, involving a flexible approach that includes foundational knowledge (Know What), practical skills for effective real-world applications (Know How), and a deep understanding of the ethical and societal implications (Know Why), enabling individuals to engage with AI technologies in a responsible, informed, ethical, and impactful manner.

AI literacy equips teachers and learners to be informed users of AI and experts in AI technologies. Essentially, it also includes an awareness of how to modify teaching and assignments to enable learners to exploit the potential of AI as a tool for learning and critical thinking.

Some practical tips for developing students' digital skills

8. Try new tools: be curious and experiment with different applications and software in your classroom. This may also inspire your students.

9. Teach critical thinking: encourage students to critically evaluate the information they see. Encourage students to encourage their peers to analyse and evaluate the information they see. Discussions about where the information comes from will help them to understand the reliability of the information.

10. Talk about security: teach students how to protect their information online. This is an important part of digital well-being.

11. Discuss ethics: raise copyright and other legal issues to help students understand the importance of responsible behaviour.

12. Encourage creativity: let students create their own digital projects. Encourage pupils to create their own projects.

¹ DigComp 2.2. (2012) JRC <https://publications.jrc.ec.europa.eu/repository/handle/JRC128415>

² Global Risks 2024: Disinformation tops global risks (2024) World Economic Forum <https://www.weforum.org/press/2024/01/global-risks-report-2024-press-release/>

³ European Parliament Think Tank (read 12.12.2024) Mis- and disinformation on social media and related risks to election integrity [https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA\(2024\)767150](https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2024)767150)

⁴ Pääkirjoitus: Britannian mellakat ovat varoittava esimerkki sosiaalisessa mediassa leviettyjen valheiden vaikutuksista. (4.8.2024) Ilfasanomat <https://www.wis.fi/paakirjoitus/art-2000010607694.html>

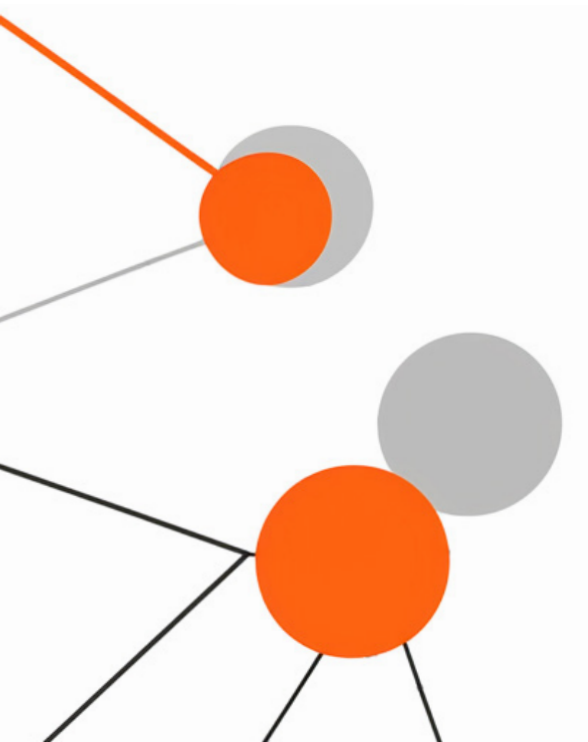
⁵ Faktabaari - www.faktabaari.fi

⁶ Nordic Observatory for Digital Media and Information disorder NORDIS (haettu 12.12.2024) <https://www.nordishub.eu/>

⁷ European Digital Media Observatory EDMO (Haettu 12.12.2024) <https://edmo.eu/>

⁸ JSN <https://jsn.fi/en/journalists-guidelines/>

- 9 Digital Information Literacy Guide (2022) Faktabaari. www.faktabaari.fi/dil
- 10 Online reading skills and strategies (2022) Digital Information Literacy Guide, Faktabaari. <https://faktabaari.fi/edu/6-online-reading-skills-strategies/>
- 11 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, <https://sciedandmisinfo.stanford.edu/>
- 12 Faktabaari www.faktabaari.fi/dil
- 13 Critical selected articles https://educritical.fi/fi/tuotokset/selected-articles/#pll_switcher
- 14 Critical <https://educritical.fi/en/>
- 15 The Framework for Digital Competence <https://eperusteet.opintopolku.fi/#/en/digiosaaminen/8706410/osaamiskokonaisuus/8706431>
- 16 PISA 2022 (2022) Valtioneuvosto. Ensituloksia lyhyesti <https://julkaisutvaltioneuvosto.fi/bistream/handle/10024/165296/Pisa22%20ensituloksia%20lyhyesti.pdf?sequence=1&isAllowed=y>
- 17 Juhani Rautopuro (2024). Mitä Pisa mittaa – ja mitä se ei todellakaan mittaa. Kieliverkosto. <https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-maaliskuu-2024/mita-pisa-mittaa-mita-pisa-ei-todellakaan-mittaa?fbclid=IwAR25limMgSGilaEPKIPuu-UvP2rzZfbzo33iDBXBTfwTA8vdQ-1961LJec8>
- 18 Hiltunen, Jenna; Ahonen, Arto; Hienonen, Ninja; Kauppinen, Heli; Kotila, Jenni; Lehtola, Piia; Leino, Kaisa; Lintuvuori, Meri; Nissinen, Kari; Puhakka, Eija; Sirén, Marjo; Vainikainen, Mari-Pauliina; Vettenranta, Jouni. 2024. PISA 2022 ensituloksia. Opetus- ja kulttuuriministeriön julkaisuja 2023:49. <https://julkaisutvaltioneuvosto.fi/handle/10024/165295>
- 19 Bozkurt, A. (2024) 'Why Generative AI Literacy, Why Now and Why it Matters in the Educational Landscape? Kings, Queens and GenAI Dragons', Open Praxis, 16(3), p. 283-290. Available at: <https://doi.org/10.55982/openpraxis.16.3.739>



6. Ethical challenges for educators in using AI

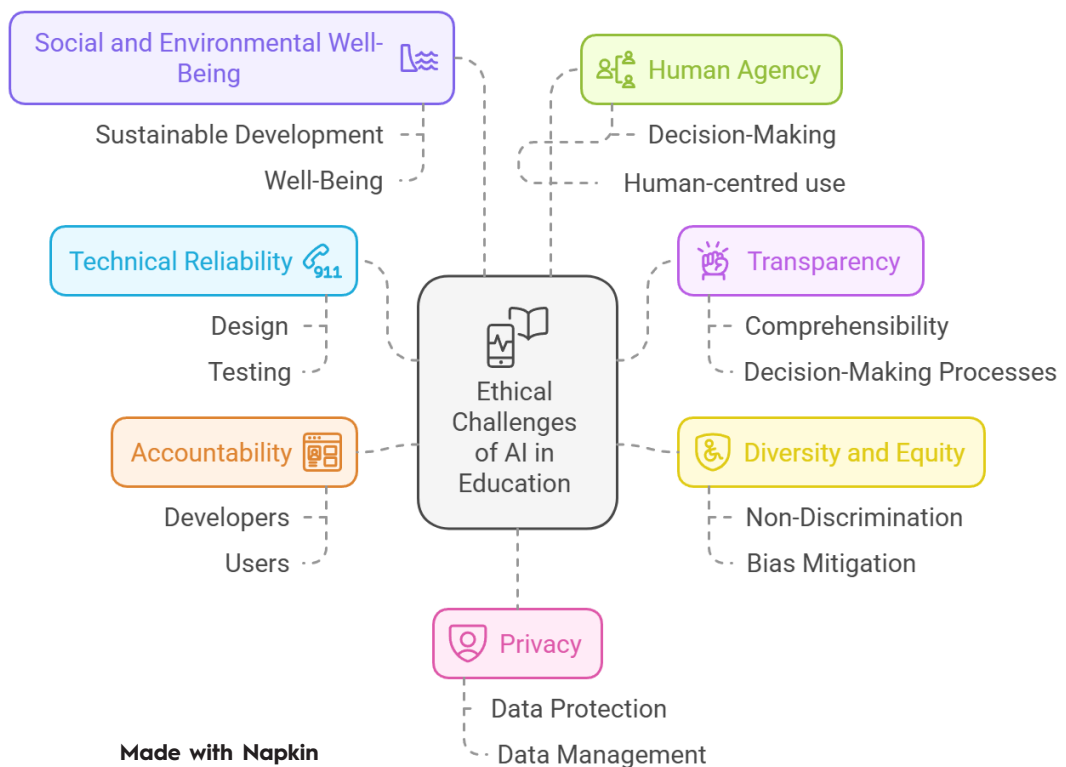
KARI KIVINEN, FAKTABAARI

THE USE OF AI IN EDUCATION requires careful planning, open communication and continuous development by teachers. This will ensure that the technology serves educational goals in a safe and pedagogical way. Teaching and assessment plans need to be adapted to ensure that the selection and use of AI tools is appropriate. New technologies should support teaching objectives and not be an additional burden.

New technology should support teaching objectives and not be an additional burden.

If properly harnessed, rapidly evolving GenAI applications have the potential to boost the efficiency of everyday school routines, create teaching materials and promote learning.

It is the responsibility of education and training providers to assess how the use of AI applications is implemented in practice, while respecting data security, data protection, intellectual property rights and the right of learners and staff to their own data and privacy. At



the same time, care must be taken to ensure that learners are not exposed to, for example, political or ideological influence, commercial advertising or marketing, or harmful or disturbing content.

Training providers must also assess whether an AI application meets the EU definition of an AI system, and with it the attached obligations and conditions.

It is particularly important to pay special attention to the ethical challenges of AI applications and the support of teachers and students, as well as the accessibility of the AI tools. The use of AI applications should support the work of teachers, improve governance and promote learning for all students in line with curriculum objectives.

Key ethical issues in the educational use of AI

The European Commission has published¹ a list of key ethical and AI-related questions for education and training providers and teachers, published by a group of experts. The questions focus on human agency, transparency, diversity, social welfare, privacy, technical reliability and accountability. The following is a summary of the ethical aspects in the use of AI in education.

Ethical requirements for education and training providers

The purchase and use of AI-based systems, tools and applications should comply with local legislation and support curricula and learning objectives. Teachers' representatives should be given the opportunity to participate in the procurement of systems and related decisions.

AI systems for educational use should be ethical, reliable, safe and human-centred. Human-centredness means, among other things, that AI systems should not be given autonomous decision-making power without human supervision, e.g. in student selection and assessment.

Laws and regulations

The EU's AI Act² classifies the use of AI in education as high-risk because it may compromise

the fundamental rights of the user, such as privacy, non-discrimination or freedom of expression. In particular, EU legislation protects the digital rights of minors.

- 1. Article 4 of the AI Act calls for measures to be taken: "Providers and deployers of AI systems shall take measures to ensure, to their best extent, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf, taking into account their technical knowledge, experience, education and training and the context the AI systems are to be used in, and considering the persons or groups of persons on whom the AI systems are to be used."**
- 2. AI systems must comply with the EU Data Protection Regulation. Access to data is restricted to those who need it, and users should know what happens to their data and how it is used. Teachers and school leaders should be aware of how to report privacy or data protection issues.**
- 3. Adequate security measures should be in place in case of a data breach, as well as control mechanisms for data processing (collection, storage, minimisation). Appropriate training should be provided for staff to comply with security guidelines.**
- 4. The potential risks of using AI tools, such as the risks of increasing commercial interests or widening the digital divide, need to be monitored and critically assessed. This will ensure that AI tools do not cause injustice or discrimination.**

Accessibility

Systems should be accessible to all learners, regardless of their gender, ability or characteristics. Particular attention should be paid to learners with special needs.

When choosing AI solutions, care should be taken to ensure that they do not widen the existing

Human agency	The development and use of AI must consider the role and impact of people. It is important to ensure that people remain at the centre of decision-making and that their rights and needs are taken into account.
Transparency	The principles and decision-making processes of AI should be clearly understood. Users and stakeholders must be able to understand how and why AI makes certain decisions.
Diversity, non-discrimination and fairness	It is important to promote diversity and prevent discrimination in the development of AI. AI systems must be fair and must not reinforce existing prejudices or discriminatory practices.
Social and environmental well-being	The use of AI must contribute to social and environmental well-being and AI solutions must support sustainable development and improve people's quality of life.
Privacy and data management	Privacy is a key issue in the use of AI. It is important to ensure that personal data is handled securely and that users have the power to control their own data.
Safety and security	AI systems must be safe to use and they should work reliably and predictably. This requires careful planning and testing to minimise the risk of errors and failures.
Responsibility and accountability	The developers and users of AI are responsible for its impact. It is important to define who is responsible for the decisions and possible consequences of AI.

Summary of the ethical aspects of the use of AI in education

digital divide. All learners should have access to the opportunities offered by AI.

It is therefore important to ensure that both teachers and students have access to the same AI tools. This will also facilitate collaboration, sharing of information and guidance, and in-service training for teachers.

Ethical expectations for teachers

AI offers new opportunities for teachers, but it also poses new ethical challenges.

Teachers should understand how AI technologies work and be aware of its potential, limitations and risks. They should also ensure in practice that AI is used ethically and to support learning.

Teachers' representatives should have the opportunity to participate in decision-making related to the acquisition and use of AI tools. Teachers should also have the opportunity to provide feedback on the tools and their performance. Staff should only use systems approved by the training provider in their teaching, in accordance with the instructions given.

Further training is needed to understand and use AI tools responsibly

Implementing AI systems requires training and some familiarity with how, for example, large language models are trained, how they work

and why there are always uncertainties in their operation that require verification of outputs.

Training providers should ensure that teachers have a good basic knowledge of the tools and software used and clear guidance on the functions, accessibility and limitations of the applications.

The provider's data protection officer should provide teachers with information on what data the systems collect about students, how the data collected is processed, what it is used for and where it is stored. Ultimately, it is the teacher's responsibility to ensure that the tools are used appropriately to support teaching and learning.

Teachers themselves should be aware of which tools have an AI component, how it is used and what the potential risks are. They also have a role to play in communicating to learners how to use the AI applications available in the school and in instructing learners how to use them safely.

AI pedagogy

AI tools should be used at pedagogically appropriate age levels and where they add value to learning. Teachers should be aware which tools are appropriate for different age groups and how they can be integrated into teaching to support learning.

Teaching should be designed to ensure that learners have equal access to AI tools, and accessibility gaps should be reported to the

education provider. This will ensure that all learners have access to AI tools, regardless of their background. AI has proven to be a very useful tool in special education, offering many possibilities for differentiation, tailoring of individualised teaching materials and tutoring to support learning.

Student assessment should never rely solely on AI tools without human verification. Teachers should ensure that AI does not replace the role of the teacher but acts as a tool and aid to support teaching.

It is important for teachers to inform learners how and why to use AI tools. This transparency will help learners understand the role of AI in teaching and reduce potential doubts, fears and misuse.

Monitoring

Artificial intelligence tools and systems are evolving at a tremendous pace. Their performance should be continuously evaluated and staff should have the opportunity to provide feedback on their experiences to training providers. This feedback will help ensure that AI tools continue to evolve and improve.

Teachers who actively participate in the use and development of AI tools can create new opportunities for learning and improve teaching outcomes. This requires ongoing dialogue, training and monitoring to ensure that AI serves, rather than threatens, educational goals.

Teachers should be given sufficient basic knowledge to be able to detect and report potential security breaches. Teachers should also be aware that AI tools can be misused, for example for academic fraud, and should be able to identify and minimise the potential risks of misuse.

Guidance

Teachers and students should be provided with clear instructions on how to use AI tools.

The Finnish National Board of Education and the Ministry of Education and Culture will publish in spring 2025 more detailed recommendations

and legislation on the use of AI in early childhood education, basic education, non-formal education, and upper secondary and vocational education and training. The aim of the support material “AI in education - legislation and recommendations” is to support education and training providers in leveraging AI as part of teaching and learning.

The support material will be published in Finnish, Swedish and English

Points to note

AI is transforming our education system by providing more personalised learning and more individualised mentoring. At the same time, it brings significant challenges.

One of the biggest concerns is the misuse of data. It is important to ensure that the personal data of learners and teachers is handled responsibly and not used as input or prompts for GenAI programmes.

Another major risk relates to the reliability of information. AI-generated information is not always completely reliable, and learners need to learn to critically evaluate AI outputs, e.g. by checking the reliability of sources.

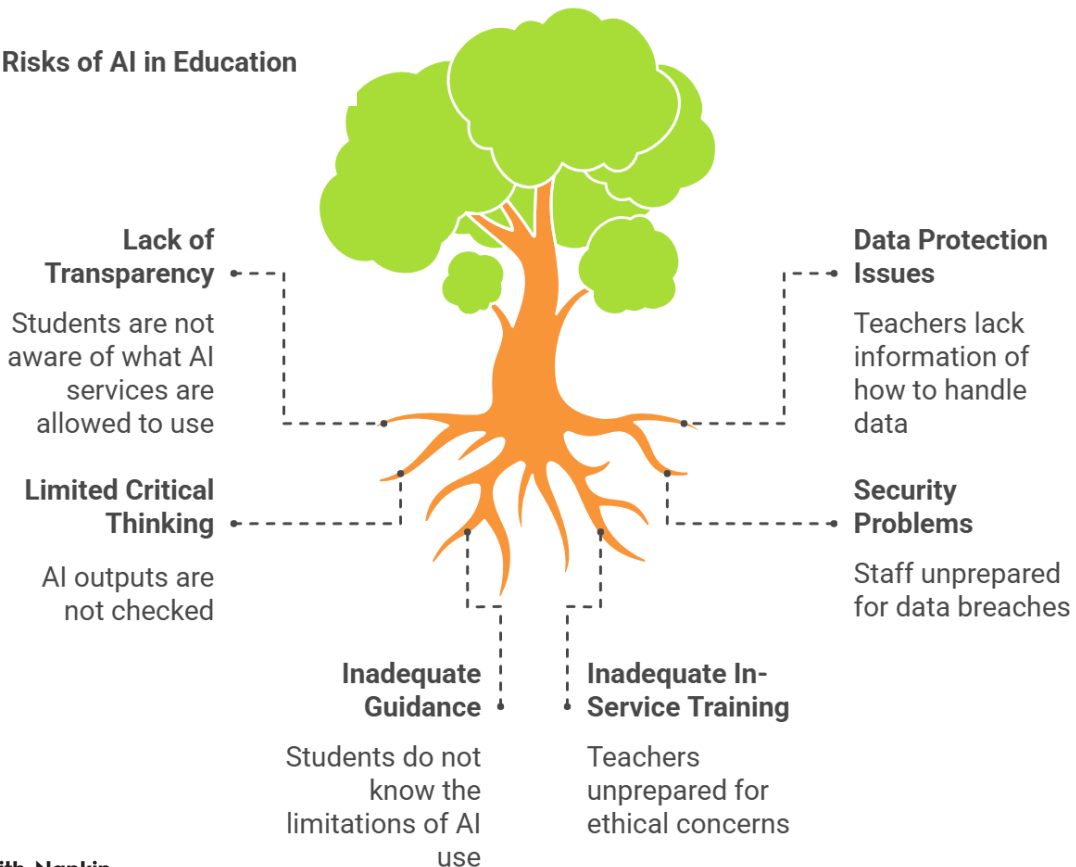
There is also a risk that learning will be overly outsourced to AI. If students rely too much on AI solutions, they will no longer develop their own problem-solving skills or delve into topics independently. This can lead to a situation where learning becomes superficial.

The commercialisation of education is also a worrying trend. AI application developers are often interested in financial gain, which can lead to situations where the content of training is tailored more to commercial than pedagogical interests.

Managing ethical risks requires teachers to be vigilant. Teachers, educational institutions and training providers need to work together to

AI systems should be accessible to all learners, regardless of their gender, abilities or profiles.

Ethical Risks of AI in Education



Made with Napkin

define clear rules for the use of AI in education. This requires ongoing discussion, training and monitoring to ensure that the use of AI serves educational objectives.

The European Commission has published “Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators”³, which contains a list of key issues for the use of AI in education, compiled by a group of experts. It is a very useful tool for developing ethical guidelines for schools and educational institutions.

Sitra’s Tarmo Toikkanen⁴ has summed up perfectly why AI outputs need human review and supervision:

Text-producing AI...

- **Does not understand the meaning of words**
- **Has no opinions, gives no valuations**
- **Does not think, does not predict**
- **Does not feel, does not care**
- **Does not take responsibility**
- **Hallucinates**

¹ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

² AI Act <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

³ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

⁴ Tarmo Toikkanen (read 31.12.2024) Tekoöly opetuksessa <https://drive.google.com/file/d/1dFhZjEWXRJn3OYP4iC-zTQJuPd5Qz-u7/view>

7. AI literacy frameworks for teachers

KARI KIVINEN, FAKTABAARI

OVER THE PAST YEAR, several AI literacy frameworks for both teachers and learners have been published.

This article briefly introduces two different frameworks for teachers. UNESCO published its own framework for both teachers and learners in autumn 2024, and a team of researchers (Filo et al., 2024) published a framework in November 2024, in collaboration with teachers and learners, which seems to be very useful.

In Finland, the Ministry of Education and the Finnish National Agency for Education are preparing recommendations on the use of AI in early childhood education and training. These will be published in spring 2025.

UNESCO's AI Competency Framework for Teachers

The UNESCO AI Competency Framework for Teachers¹ was launched in September 2024. The framework focuses on teachers' lifelong professional development and includes five key points.

The framework aims to ensure that teachers are equipped to use AI responsibly and effectively, while minimising potential risks to students and society.

- **The human-centred approach defines the values and critical attitudes that teachers need to develop towards human-AI interactions. Teachers are encouraged to always place human rights and human welfare needs at the centre of AI**

Aspects	Progression		
	Acquire	Deepen	Create
Human-centred mindset	Human agency	Human accountability	Social responsibility
Ethics of AI	Ethical principles	Safe and responsible use	Co-creating ethical rules
AI foundations and applications	Basic AI techniques and applications	Application skills	Creating with AI
AI pedagogy	AI-assisted teaching	AI pedagogy integration	AI-enhanced pedagogical transformation
AI for professional development	AI enabling lifelong professional learning	AI to enhance organisational learning	AI to support professional transformation

in education. Teachers are encouraged to develop critical approaches to assessing the benefits and risks of AI, to care about human agency and responsibility, and to understand the social and civic implications of AI in the AI era.

- **Ethics of AI:** Teachers are aware of the basic ethical principles related to AI and understand their importance, recognise the human-centred nature of AI and the central role of people in all stages and aspects of AI development. AI ethics includes the key ethical values, principles, regulations, institutional laws and codes of practice that are important for teachers to understand and apply, drawn from the rapidly expanding body of knowledge on AI ethics and its implications for education.
- **The foundations and applications of AI** defines the basic knowledge and practical skills that teachers need to support the selection, application and creative tailoring of AI tools to build student-centred AI-assisted teaching and learning environments. Teachers are expected to understand the definition of AI, basic knowledge of how AI works and the main categories of AI technologies. They should also have the necessary skills to assess the usefulness and limitations of AI tools.
- **AI pedagogy:** teachers can identify and assess the pedagogical benefits of AI applications and choose appropriate tools to support instructional design, pedagogy and assessment. Teachers need to develop the ability to critically assess when and how AI can be used in teaching and learning in an ethical and human-centred way, and to design and implement inclusive AI-assisted teaching and learning practices.
- **AI for professional development:** teachers are aware of the potential of AI to support their continuous professional development and are motivated to use AI for professional lifelong learning. Given the rapid

development of AI, teachers need guidance on how to continue their professional development in educational environments with increasing interaction between humans and AI.

The aspects of the framework are interlinked and complementary, not separate. Effective teaching (with or without AI) requires a holistic approach that combines different areas of expertise.

For example, a teacher's ability to apply AI pedagogy is influenced by their understanding of the basics of AI, their awareness of AI guidelines and their commitment to continuous professional development. Similarly, a teacher's ability to cope with the ethical problems of AI is influenced by their understanding of the basics of AI and their experience of applying AI in teaching. Skills in one area can enhance skills in another. A deeper understanding of the fundamentals of AI can improve a teacher's ability to apply the pedagogical and ethical principles of AI, and continuous professional development will improve a teacher's understanding of all these aspects.

The framework stresses that AI tools should complement, not replace, the vital roles and responsibilities of teachers in education.

The second and relatively new framework for teachers' and students' AI competences, developed with teachers and students, defines four key competences²:

- 1. Identification of AI mechanisms and their operation:** An understanding of the principles of how AI works helps to assess the potential and limitations of its use.
- 2. Effective and informed use of AI:** The ability to create effective inputs and to critically review the results.
- 3. AI Agency, proactive and value-generating utilisation of AI:** Using AI responsibly and creatively to create added value.
- 4. Ethical use:** The use of AI should be based on ethical principles.

The framework also outlines the values, attitudes and knowledge needed to deal with AI in education, to help prepare teachers and students for the AI-saturated world.

According to the proposed framework, identifying AI mechanisms is a critical skill for both teachers and students, as it lays the foundation for navigating and critically processing AI applications in different contexts.

Understanding the principles of how AI works will enable students to better assess the potential, limitations and ethical implications of AI applications in everyday life.

Effective and informed use of AI involves understanding how to create a feed to achieve the desired results and being able to identify the right AI tool for a given task. The third component involves critically examining the results produced by AI tools, assessing their accuracy, identifying potential biases, and understanding the ethical implications of their use in everyday life.

AI Agency means understanding that ethical and fair use of AI means to empower people without replacing their contributions. This approach requires a deep understanding of the potential and limitations of AI. AI agency involves the active and creative use of AI to empower people, not just its passive use. This emphasises the creative application of AI in problem solving and value

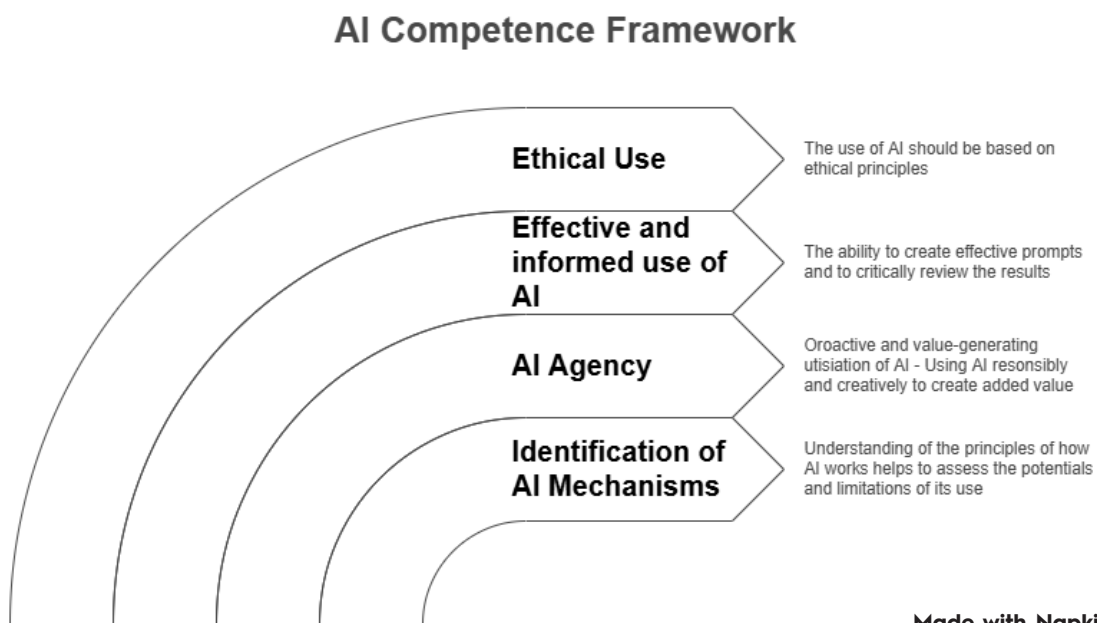
creation, which is a major advantage in today's world. This will enable people to use AI in a positive way to impact their lives and the world around them. At the same time, it emphasises the need to understand human-machine collaboration and to critically evaluate the results of AI.

The ethical use of AI aims to help learners make good decisions. Fairness, privacy, morality and responsible decision-making should be emphasised to learners in the use and design of AI tools. Ethical use of AI means identifying and correcting biases and distortions in AI systems, reducing inequalities, and promoting important human values and the well-being of society. Ethics also requires careful assessment of the long-term effects of AI, such as the degradation of human capabilities, discrimination, or the manipulation of human perceptions. At the same time, it emphasises the responsibility and active role of individuals in preventing these challenge

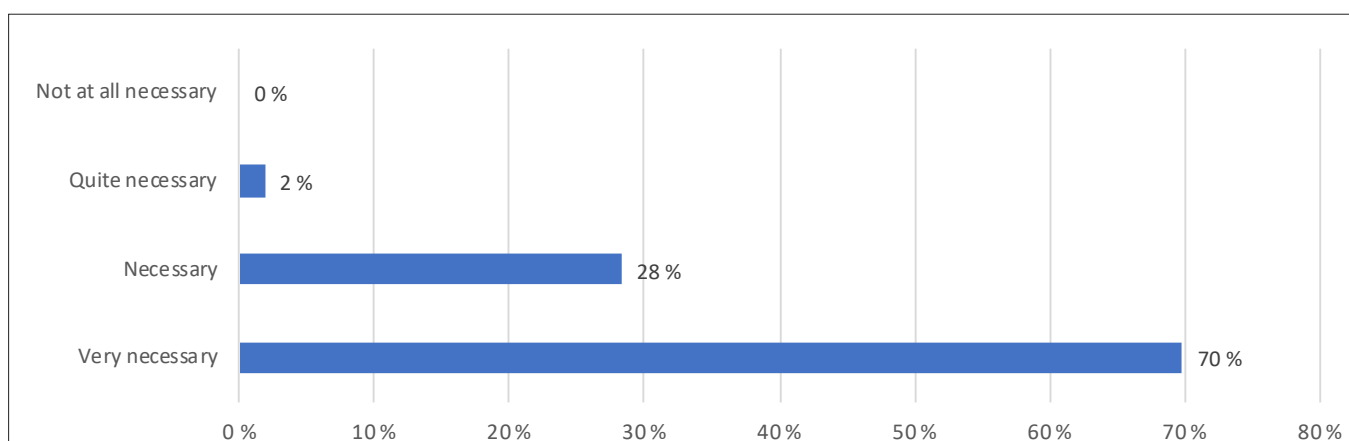
Finnish AI recommendations

In October 2024, the Ministry of Education and Culture and the Finnish National Agency for Education issued the first part of the AI recommendations for consultation. The following key criteria have been extracted from it:

Understanding how AI works and its implications is a prerequisite for the responsible and safe



How necessary do you think the Finnish National Board of Education's recommendations on the use of AI are?



use of AI. This requires a critical assessment of the potential and limitations of AI and a holistic consideration of the benefits and drawbacks.

Before implementing AI in education, it is necessary to ensure that users have sufficient understanding and ability to interpret and critically evaluate the information generated by AI.

A key challenge for education and training is to equip learners with skills such as critical thinking, problem solving, source criticism and creativity in the use of AI.

In a digitalised and globalised world, these skills are a prerequisite for active participation and for ensuring freedom of expression and democracy.

The ethics of AI and the ethical use of technology will become a key issue in education and training.

The final formulations of the recommendations and supporting materials will be published in 2025 to support early childhood education and care, teaching and training and learning.

At the Finnish Regional State Administrative Agency (AVI) AI training in October 2024, teachers who attended were asked how necessary they thought these recommendations were. The answer was clear: 98% of teachers answered that the recommendations were necessary or very necessary.

¹ UNESCO AI Competency Framework for Teachers (2024) <https://unesdoc.unesco.org/ark:/48223/pf0000391104>

² Filo, Y., Rabin, E. & Mor, Y. An Artificial Intelligence Competency Framework for Teachers and Students: Co-created With Teachers. *European Journal of Open, Distance and E-Learning*, 2024, Sciendo, vol. 26 no. s1, pp. 93-106. <https://doi.org/10.2478/eurodl-2024-0012>

8. AI assessment scales

KARI KIVINEN, FAKTABAARI

MANY TEACHERS HAVE expressed concern that students may use GenAI tools for their homework, which can be seen as cheating, plagiarism or even academic fraud. Therefore, learners should be encouraged to be open about the use of AI in their work.

When including AI content in research, articles or teaching, it would be advisable to cite the AI contribution in the same way as one cites other sources. For example, the University of Jyväskylä¹ provides the following documentation guidelines:

“When using AI applications in your studies or scientific work, document your use carefully. Documentation is a prerequisite for being able to report your use of AI truthfully and in accordance with good scientific practice.

According to the University of Jyväskylä’s policy, students must report the use of generative AI applications in their learning tasks and theses.”

Generative AI services are available free of charge to all learners. Therefore, teachers should consider what kind of assessment tasks they give and what they want to assess. It is no longer appropriate to simply give tasks that require checking information or writing essays. Search engines and GenAI make it easy for learners to search for any information and produce texts on any topic.

Teachers should discuss with students beforehand how AI can be used for the exercise to be assessed. The AI Assessment Scale (AIAS) proposed by Furze & AI² aims to clarify communication between teachers and students. The assessment framework can be used to make it clear to both students and staff what is appropriate or inappropriate use of AI in an assignment.

For example, the use of AI can be banned in some important exams. On the other hand, there are many tasks where AI can be used as a useful brainstorming or support tool. Creative use of AI may be favoured for tasks requiring problem solving and innovation.

If restrictions on the use of AI are to be enforced, they must be clearly stated, and clear guidance should be available to all. The updated Furze table also includes clear instructions for learners in different situations.

Guidance for learners on the use of AI

If we want to guarantee a fully AI-free assessment of learners' knowledge and skills, we need to have some form of supervised assessment. According to Furze³, this is for two reasons: "First, there is no way to guarantee that a student who has a device (mobile device, laptop, tablet, Meta Ray Bans...) will not be AI-assisted. Second, if we try to use recognition tools or other technology solutions, we risk creating equity issues between, for example, students who have access to (and get away with) more advanced, paid AI products and students who have limited access to (and are likely to get away with) free tools."

Many educational institutions have implemented AI detection tools to monitor the use of AI. Unfortunately, they have proven to be unreliable and often produce false positives⁴. Reliance on these tools can lead to unfair accusations against innocent students, but at the same time may miss sophisticated misuse by students who deliberately seek to avoid detection.

If students are given permission to use AI, it is advisable to ask them to document their working process; for example, the prompts/inputs they used, the results they obtained, and how they worked on and modified the results they obtained. This method not only prevents inappropriate use, but also helps students develop key AI skills. In addition, seeing the work in progress gives teachers some idea of the potential applications of these tools in the tasks they are assigned⁵.

¹ Jyväskylän yliopisto (read 15.12.2024) <https://www.jyu.fi/fo/opiskelijalle/kandi-ja-maisteriopiskelijan-ohjeet/tiedonhankinta-ja-aineistonhallinta/tekoaly-tiedonhankinnassa>

² The updated AI assessment scale (noudettu 12.12.2024) <https://leonfurze.com/2024/08/28/updated-ai-assessment-scale/>

³ Leon Furze (haettu 29.12.2024) <https://leonfurze.com/2024/08/09/can-the-ai-assessment-scale-stop-students-cheating-with-ai/>

⁴ Jason M Lodge (read 29.12.2024) The evolving risk to academic integrity posed by generative artificial intelligence: Options for immediate action <https://www.teqsa.gov.au/sites/default/files/2024-08/evolving-risk-to-academic-integrity-posed-by-generative-artificial-intelligence.pdf>

⁵ Cf. Jason M Lodge (read 29.12.2024) The evolving risk to academic integrity posed by generative artificial intelligence: Options for immediate action <https://www.teqsa.gov.au/sites/default/files/2024-08/evolving-risk-to-academic-integrity-posed-by-generative-artificial-intelligence.pdf>

Guidance for teachers on the use of AI

1	No AI	The assessment is completed entirely without AI assistance in a controlled environment, ensuring that students rely solely on their existing knowledge, understanding and skills.
2	AI planning	AI may be used for pre-task activities such as brainstorming, outlining and initial research. The level focuses on the effective use of AI for planning, synthesis, and ideation, but assessments should emphasise the ability to develop and refine these ideas independently.
3	AI collaboration	AI may be used to help complete the task, including idea generation, drafting, feedback and refinement. Students should critically evaluate and modify the AI suggested outputs, demonstrating their understanding.
4	Full AI	AI may be used to complete any elements of the task, with students directing AI to achieve the assessment goals. Assessments of this level may also require engagement with AI to achieve goals and solve problems.

Leon Furze et al (2024) updated AIAS scale

Guidance for learners on the use of AI

1	NO AI	You must not use AI at any point during the assessment. You must demonstrate your core skills and knowledge.
2	AI planning	You may use AI for planning, idea development and research. Your final submission should show how you have developed and refined these ideas.
3	AI collaboration	You may use AI to assist with specific tasks such as drafting text, refining and evaluating your work. You must critically evaluate and modify AI-generated content you use.
4	Full AI	You may use AI extensively throughout your work either as you wish, or as specifically directed in your assessment. Focus on directing AI to achieve your goals while demonstrating your critical thinking.
5	AI exploration	You should use AI creatively to solve the task, potentially co-designing new approaches with your instructor.

Leon Furze et al (2024) updated AIAS scale

9. What should a teacher know about the EU AI Act?

TARMO TOIKKANEN, SITRA

“The deployment of AI systems in education is important to promote high-quality digital education and training and to allow all learners and teachers to acquire and share the necessary digital skills and competences, including media literacy, and critical thinking, to take an active part in the economy, society, and in democratic processes.”

– AI ACT preface (56)¹

AI SOLUTIONS CAN BE valuable in education, but they need to be developed and deployed responsibly. If we simply let the first AI chatbot we find assess students’ work and then record it in the student’s records, we are being far from responsible. To make it clear to everyone what responsible means, the EU AI Act sets out the situations in which we need to take particular care.

Organisations are obliged to ensure their employees are AI literate.

The responsibility for AI solutions in education lies with the institution, not the teacher. On the other hand, all organisations are obliged to ensure their employees are AI literate. The educational institution must therefore train its teachers to work responsibly with AI systems. In this short article, I will describe the general features of the AI Act to help teachers understand the big picture.

The main message, however, is that the educational institution must provide teachers with adequate guidance on how to work with different types of AI systems.

Towards reliable AI services

AI can be added to almost anything: cars, toys, recruitment, law enforcement or even student assessment. It’s clear that poorly designed AI in, say, a self-driving car or the grading of students’ essays can cause all sorts of trouble.

There are two kinds of harm. It can:

1. pose risks to people’s fundamental rights, or
2. it can lead to dangerous products.

The aim of the AI Act is to ensure that only trustworthy AI services are available on the EU market.

Fundamental rights include the right to education, the right to work, freedom of expression, freedom of religion and equality. Dangerous products are regulated through product regulation. The EU has laws regulating the safety of cars, lifts, food and electrical appliances, for example. The CE marking shows that the manufacturer promises that the product is safe. Similar thinking is being extended to the AI part of these products.

The aim of the AI Act is to ensure that only trustworthy AI services are available on the EU market. Conversely, this means that if AI can cause harm (or risks), it must be made trustworthy – that is, it must be designed properly.

Doing it right takes a lot of work, but it can be summed up by one thing: good data governance. When data flows are identified, described and managed, it also becomes easier to ensure accountability and reliability. Data governance is the solution to both data protection issues and ensuring the trustworthiness of AI. Quality control and risk management should be integrated into the same exercise.

AI systems can be classified according to their risks

- 1. Unacceptable risk: some extreme use cases that are not in line with European values are banned altogether. These include social scoring of citizens by combining different databases, real-time biometric identification of citizens or cognitive-behavioural manipulation of people.**
- 2. High risk: AI is part of a regulated product or may otherwise pose a risk to fundamental rights. This is permissible, but AI must be properly built, deployed, documented and approved by the authorities.**

- 3. Limited risk: for example, chatbots, image generators and the like that produce content for human viewing. The only obligation is to inform the user that they are dealing with AI.**
- 4. No risk: certain background processes, such as spam filters or AI-generated terrains in video games. There are no obligations for these.**

In the education sector, high-risk AI systems include:

- **Student selection**
- **Assessment of learning outcomes**
- **Assessing a person's future educational needs**
- **Monitoring students during exams**

The above list is indicative, and the level of risk is influenced by how autonomously the AI system operates. If, for example, the AI only assists the teacher in assessing learning outcomes, it may not be high risk. If, on the other hand, it autonomously generates grading recommendations, the level of risk is higher.

How to acquire AI

An AI system provider needs to know the risk level of its service. The CE label on an AI system indicates that it has been manufactured in accordance with EU quality standards.

The AI Act requires certain reliability assurance measures when deploying high-risk AI. However, the same measures are good practices that should be followed even if they are not mandatory.

While the AI vendor has done their part and shared them with you, as a deployer you must do your part.

- **Data management: how the system is used, what data can be entered into it and who has access to it. Data confidentiality, data protection issues and ethical considerations will be considered.**

- **Quality control:** How the system's performance is monitored and how errors are detected. Monitoring and description by professionals is of paramount importance.
- **Risk management:** how to recover from incidents, security problems, etc.

How to develop AI

As an AI system developer, you need to identify your customers' use cases and their highest level of risk. If it's a high-risk system, AI regulation will necessitate several steps, but it's worth it.

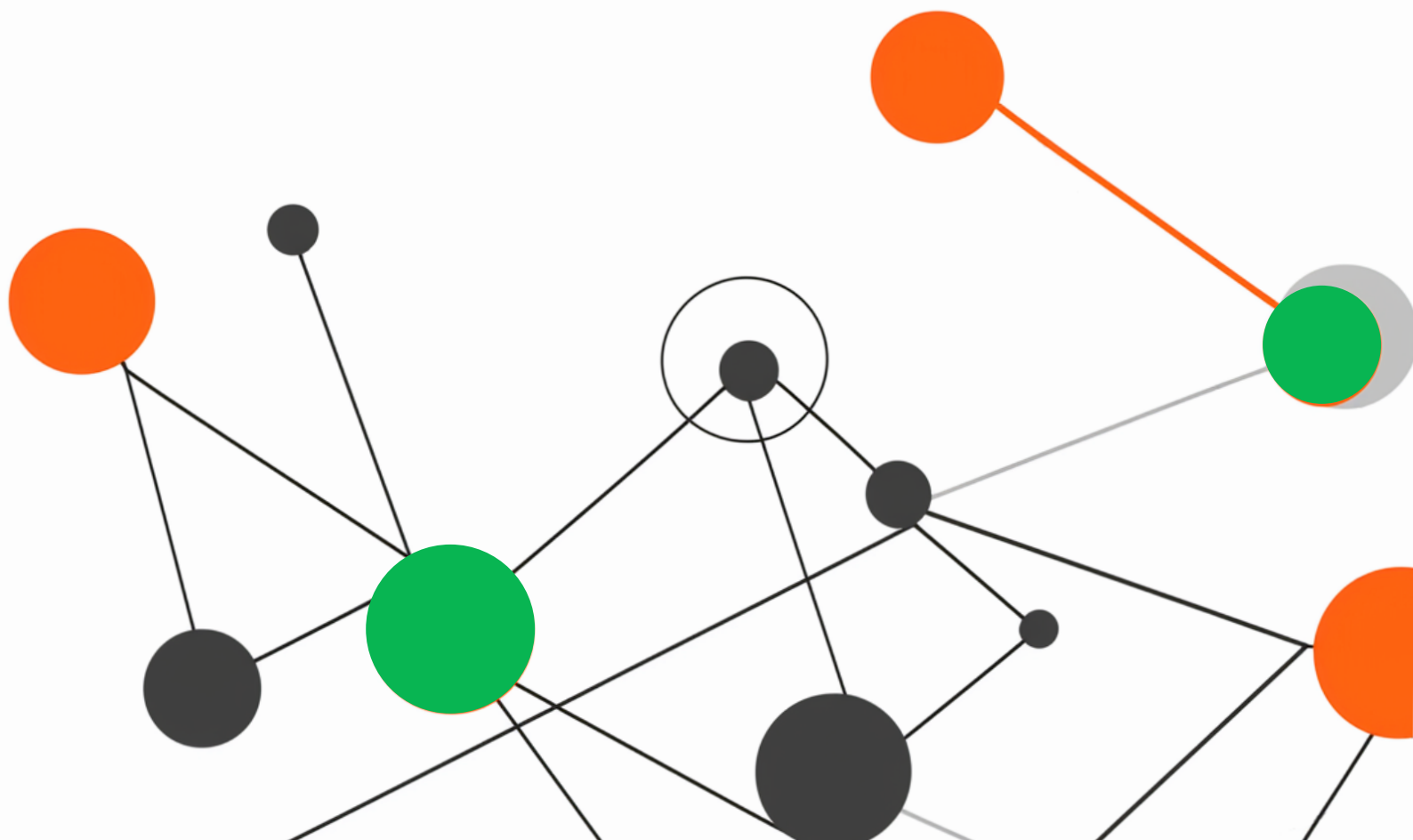
- **Data management:** Data sources, metadata, training algorithms, data access.
- **Quality control:** Automated test suites, training-quality monitoring, in-service monitoring.
- **Risk management:** Systemic risks, developer's own risks, risks caused by customers' activities. Risk management plan.

Additional information

Traficom is the single point of contact for the Finnish supervisory authorities, and its website provides more detailed guidance and contact details for the different authorities.

Sitra is producing a self-study course on AI regulation as part of the ABC of Data Economy course series. This course will be ready in spring 2025 and published at <https://www.sitra.fi/en/projects/basics-of-the-data-economy/>

¹ AI ACT Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)Text with EEA relevance.



10. Teaching AI skills through “Generation AI” applications

JARI LARU, DR, SENIOR LECTURER, FACULTY OF EDUCATION AND PSYCHOLOGY, UNIVERSITY OF OULU.

GENERATION AI IS a project funded by the Strategic Research Council (STN) of the Academy of Finland.

The overall aim of the Generation AI project is to support children and young people in understanding AI and machine learning technologies. It also aims to help them critically and ethically examine the different backgrounds, uses and impacts of AI-based technologies. Another important starting point is to support children's creative capacities and their growth as informed and responsible actors for the future. The aim is for pupils to have meaningful experiences of success and to feel empowered to make a difference to the things they care about.

We will develop applications to support AI education that are safe, accessible and free.

We will develop applications that are all safe, responsive and accessible. We will not collect data or track you. The legal experts working on our project have ensured that our apps comply with EU data protection requirements. In addition, our apps will be free and released under an MIT licence.

Teachable machine

AI applications have a strong presence in our everyday lives. We encounter them on social media, in innovative text and image generators, or even when we unlock our phones using facial recognition. But few of us think about how AI works, how AI applications are developed, and why AI sometimes makes mistakes.

Teachable machine is a tool that allows students to learn how AI works by building simple machine vision applications based on classification. It allows students to work through the entire machine learning workflow, from collecting data to deploying a mobile app as part of their work, without any programming skills. The chosen approach reduces the need for technical skills, making AI learning more inclusive¹.

The app is suitable for all ages and has been specifically designed with children and young people in mind. The teachable machine illustrates the potential of AI in a concrete way, introduces the underlying concepts and also leads to reflection on the limitations and ethical issues of AI.

Social media machine - social media simulator

The social media machine is an explainable artificial intelligence (XAI) tool designed for primary school students. It provides a familiar, interactive interface that matches students' experiences on popular social media platforms, while offering ways to "peek under the hood"

and reveal how "instagram works". The aim is to teach key data-driven and AI-based social media mechanisms and core concepts such as data collection, profiling and recommendation in an accessible and engaging way.

The aim is also to develop students' understanding of data literacy, AI literacy and AI ethics².

Through the social media machine, students will gain first-hand experience of how even the smallest actions, such as stopping content viewing, are recorded in their digital footprint and further in their personal profile. The social media machine uses real-time visualisations that allow users to see how their actions and those of others affect their social media experience³.

The social media machine learning materials and more detailed user manuals were published in

the 2025 EDUCA event and are now available on the project website.

Involving children, young people and teachers as active participants and developing applications and materials together

Generation AI organises an annual participatory co-design project (school project) for participating schools (schools in Joensuu and Oulu), which includes age-appropriate, hands-on activities and interdisciplinary research together with teachers and students. The project is designed so that the same children and young people participate in three school projects (see Table 1) in successive years, introducing basic concepts and skills and addressing the social implications of everyday life. As part of the joint design, learning materials⁴ and tools⁵ will also be developed to become national capital for every teacher and pupil in our country.

The first school project was carried out in 2023 and included data-driven design. It involved school students designing machine learning-based classifier applications using a teaching machine developed in the project⁶. In practice, the students participated in three workshops on the basic concepts and mechanisms of AI, data-driven design and the impact and ethics of AI⁷.

The second school project in 2024 dealt with social media mechanisms, which were studied in two different lessons. In the first lesson, we studied data collection, profiling and recommendation using a machine learning exercise and a profiling game. In the second lesson, the social media machine simulator⁸ was used to study how social media mechanisms actually work. Materials related

Materials on children's AI projects will be published during spring 2025

Teachable machine:

<https://tm.generation-ai-stn.fi>

The source code of teachable machine:

<https://github.com/knicos/genai-tm>

Teaching materials:

<https://www.generation-ai-stn.fi/materiaalit>

Social media machine:

<https://someone.generation-ai-stn.fi>

The source code of the social media machine:

<https://github.com/knicos/genai-somekone>

Social media machine teaching materials:

<https://www.generation-ai-stn.fi/materiaalit>

School project	Name	Concepts	Lessons	Tool
2023	Data-driven design (how AI works)	Educational data, classifier, certainty, fragility, algorithmic bias	Basic concepts and mechanisms of artificial intelligence Data-driven design Impacts and ethics of AI	Teachable machine + materials
2024	Mechanisms of social media	Data collection, profiling, recommendation	Basics and profiling game How does social media work? The effects of Social media use?	Profile game, Social media machine + materials,
2025	The impact of social media	Social media bubbles, unintended effects, shaping behaviour and opinions	Under development	Under development

School projects 2023-2025

Project	Lesson	Task description
Robotit (esikoulu)	1	• Children draw robots on paper and on an interactive whiteboard
	2	• Kids build robots from recycled materials
	3	• Teacher-led story development using AI • Children present the robots they have built
Robottitehdas (peruskoulu, 1 luokka, erityiskoulu)	1	• Children draw robots
	2	• Teacher has used AI to create images to be studied • Children draw, design, draw and build robot factories • Teacher uses generative AI to make images from children's designs
	3	• Children develop stories about their own robots in groups of three. • Teacher creates a collective story on ChatGPT based on the children's descriptions
Super heros (pre-school)	1	• Children draw superheroes with wood colours • Children tell what superhero characteristics they have
	2	• Teacher uses generative AI to make pictures based on children's drawings
	3	• Teacher creates a story based on children's descriptions using generative AI

Example tasks for pre-primary and primary education

Step	Title	Description
1	Examine the profiles created when using a social networking site	<ul style="list-style-type: none"> • How are images and keywords related? • What content is important/interesting to these people? • What kind of visual content appeals to these people and attracts their attention?
2	Design an advertising campaign for this user group	<ul style="list-style-type: none"> • What is the backstory and message of the campaign? • What in particular do you want to highlight about your chosen topic (originally from Helsinki) • What is the campaign slogan?
3	Use generative AI to create the promotional material needed for your campaign	<ul style="list-style-type: none"> • Images • Leaflets, stories or other texts needed for your ads. • Videos

Phases in the campaign "Helsinki, the summer city of youth"

to this were published in January of 2025! Similarly, the planning for the third school project is just starting, so more information will be available later.

4. How to use AI creatively - towards the ability to use AI creatively and responsibly

Epistemic doubt - this rather awkward-sounding term has been the starting point for thinking about how we can learn to live, be, work and learn in this world of AI technologies. With the advent of generative AI, each of us has some doubts about the reliability of the content produced by AI. Total distrust prevents us from working with AI technologies, so it is a kind of balancing act⁹.

Pre-primary and primary education

In the context of pre-primary and primary education, research to date has looked at what can be done with AI and how content can be integrated into the pre-primary curriculum. This has led to projects such as Robotics, Robot Factory and Superhero, all of which have used generative AI as part of the pre-primary curriculum. AI has been used for storytelling, image-making and questioning, among other things.

From the above projects, researchers have identified different metaphors related to AI used by children and teachers: a) what AI is: computer (program); human (assistant); fantasy (magician); b) how AI works: computer (commands); human (intelligence, action); comparing human and technology; comparing natural and digital world; comparing fantasy and human action; b) existence of AI: doubt/belief; reassurance.

Basic education

In the context of basic education, various projects related to the use of generative AI have been carried out together with students in basic education, including projects related to luxury art and the internet, media and marketing.

The table below (Table 4) describes an example task in which the 6th grade students were asked to

create advertisements for the advertising campaign "Helsinki, the summer city of youth". In this fictional project, the 6th grade students used a soma machine to create different youth profiles and then were tasked with using generative AI to create targeted marketing materials for young people.

Our applications and materials will take into account children's digital rights, as we will provide information on children's digital rights and EU regulation.

In addition to developing tools, methods and teaching materials, the Generation AI project provides Finnish teachers and education providers with information on national legislation and EU regulations. These include, for example, data protection and AI regulation. Professor Susanna Lindroos-Hovinheimo has published several articles on the EU AI Regulation, including a short introduction to the EU AI Regulation published in the Yearbook of Communication Law last year¹⁰.

Digital environments are often not designed with children or children's rights in mind, but the tools have wide-ranging implications for children's rights. There is also a clear tension identified between, for example, the need to protect children and children's autonomy. For example, how can we protect children from risks, but at the same time enable children to access digital applications in the school environment? On this topic, Lakimies magazine¹¹ published an openly available research article in Finnish, which has already received well-deserved attention, e.g. from education and training providers.

Teachers and other adults have a key role to play in ensuring children's rights and their practical implementation - it is not enough to talk about them only at the top. The classroom and everyday school life is a particularly crucial context for the realisation of children's digital rights. The Generation AI Privacy Guide¹², published on the project's website, can help.

Materials from AI projects in the context of basic education will be published during spring 2025

Summary

The applications, methods, materials and guides developed in Generation Ai are based on peer-reviewed and high quality international scientific research. We will build on this to develop the basis for AI and safety education in our country.

There is a strong demand for this project, as e.g. research on children's rights is scarce in the context of AI and pedagogical models for research-based teaching of AI and data literacy are scarce.

We are developing solutions that are rare even on an international level, so you, the Finnish teacher, are privileged!

Follow our journey and join us!

-
- 1 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024a). A no code AI education tool for learning AI in K-12 by making machine learning-driven apps. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
 - 2 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024b). An educational tool for learning about social media tracking, profiling, and recommendation. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
 - 3 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024b). An educational tool for learning about social media tracking, profiling, and recommendation. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).
 - 4 Kahila, J., Vartiainen, H., Tedre, M., Laru, J., Arkko, E., Lin, A., Pope, N. & Jormanainen, I. (2024a). Tekoäly oppimisen kohteena ja luovan toiminnan lähteenä. Valkoinen kirja. January Collective. <https://www.generation-ai-stn.fi/kirjasto/opetettavakone/mika-on-opetettava-kone/opetettava-kone-ja-materiaalit-perustuvat-tutkimukseen/>

Kahila, J., Vartiainen, H., Tedre, M., Arkko, E., Lin, A., Pope, N., Jormanainen, I. & Valttonen, T. (2024b). Pedagogical framework for cultivating children's data agency and creative abilities in the age of AI. Informatics in Education. doi:10.15388/infedu.2024.15

5 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024a). A no code AI education tool for learning AI in K-12 by making machine learning-driven apps. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT). Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024b). An educational tool for learning about social media tracking, profiling, and recommendation. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).

6 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024a). A no code AI education tool for learning AI in K-12 by making machine learning-driven apps. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).

7 Kahila, J., Vartiainen, H., Tedre, M., Laru, J., Arkko, E., Lin, A., Pope, N. & Jormanainen, I. (2024a). Tekoäly oppimisen kohteena ja luovan toiminnan lähteenä. Valkoinen kirja. January Collective. <https://www.generation-ai-stn.fi/kirjasto/opetettavakone/mika-on-opetettava-kone/opetettava-kone-ja-materiaalit-perustuvat-tutkimukseen/>

Kahila, J., Vartiainen, H., Tedre, M., Arkko, E., Lin, A., Pope, N., Jormanainen, I. & Valttonen, T. (2024b). Pedagogical framework for cultivating children's data agency and creative abilities in the age of AI. Informatics in Education. doi:10.15388/infedu.2024.15

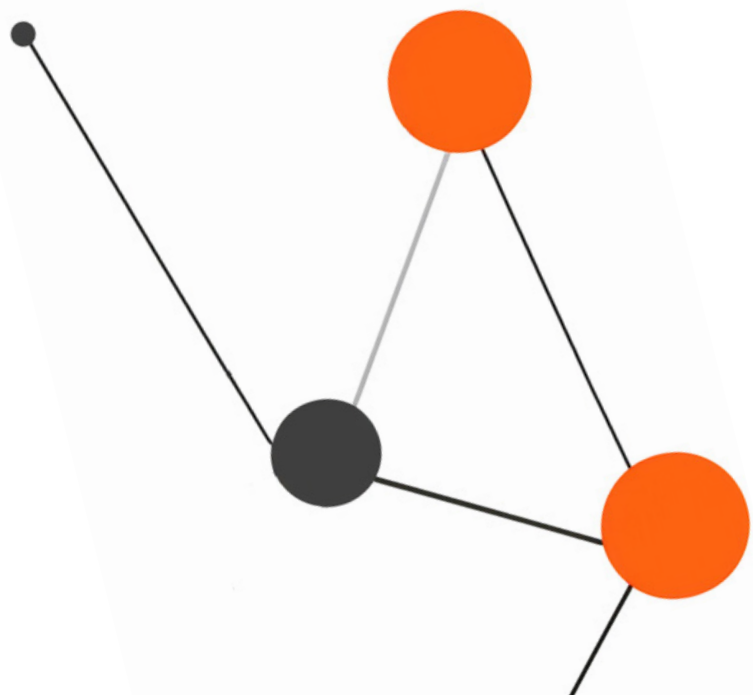
8 Pope, N., Vartiainen, H., Kahila, J., Laru, J. & Tedre, M. (2024b). An educational tool for learning about social media tracking, profiling, and recommendation. 2024 IEEE International Conference on Advanced Learning Technologies (ICALT).

9 Davies, S., Sormunen, K., & Kangas, K. (2024). Cultivating Epistemic Doubt: a Key Competence for Productive Participation in the Era of Artificial Intelligence. 17th International Conference on Computer-Supported Collaborative Learning (CSCL) 2024 - Buffalo, Yhdysvallat (USA)

10 Lindroos-Hovinheimo, S. (2023). Lyhyt johdatus Unionin tekoälyasetukseen. In Viestintäoikeuden vuosikirja 2023 (pp. 67-76). Forum Juris, Helsingin yliopisto Oikeustieteellinen tiedekunta. <https://www.edilex.fi/viestintaoikeuden-vuosikirja/1000880004>

11 Silvennoinen, E., Tedre, M., & Valttonen, T. (2024). Datafikoitua peruskoulu-tasapainoilua lapsen henkilötietojen suojan ja opetuksen tavoitteiden välillä. Lakimies, 122(5), 655-678.

12 Generation AI Tietosujoaopas (Haettu 18.12.2024) <https://www.generation-ai-stn.fi/henkilotietojen-kaasittely-alakoulussa/>



11. Finnish “Generation AI”

JARI LARU, DR, SENIOR LECTURER, FACULTY OF EDUCATION AND PSYCHOLOGY, UNIVERSITY OF OULU.

GENERATION AI IS one of the projects of the Shield programme

Generation AI is one of the projects of the Security and Trust in the Age of Algorithms (Shield) programme. Shield is a set of AI projects funded by the Strategic Research Council (STN) of the Academy of Finland.

The Shield programme focuses on security and trust in a world of information technology and hybrid influencing. In addition to more traditional security threats, threats and opportunities created by digitalisation and technological development as well as different crises from biodiversity loss to climate change, financial crises and power-political conflicts are shaping our operating environment in unpredictable ways.

Why is the Generation AI project important right now?

Generation AI addresses three interlinked societal challenges

1. **The deterioration of security, control and a sense of free will in a data-driven society**
2. **the erosion of trust in public authorities, the media, science, civil society and citizens; and**
3. **rapidly growing inequalities in the management and empowerment of ever-changing ICTs.**

The project will develop interventions in the context of pre-school and school education to increase children’s resilience in a technologically evolving world, to create and promote a cybersecurity mindset and to equip teachers with the insights and pedagogies needed to change curricula. The project will also enhance research on children’s rights (CRC) in the field of AI.

What will the project do?

The project will build a foundation on which to build technology education for the “AI generation”: technology education that takes into account how AI works, the opportunities it offers and its dynamic effects - without ignoring its shortcomings and risks, or how AI systems

can foster confrontation, discrimination, insecurity and erode trust.

The project will support learners' understanding of safety and trust in AI and machine learning (ML)-based systems. The project will measure its impact on learners' data literacy and their ability to work with complex and non-linear processes, making them less unpredictable and polarising.

The project will enable learners to recognise the consequences of the often unconscious choices they make online. It illustrates how biases creep into systems and create algorithmic discrimination and inequalities. It sensitises learners to questions of governance, technological determinism, and who has power over data and in what ways users' trust can be abused.

Who is the target group of the Generation Ai project?

The project brings together technology developers, schools, public authorities, businesses and NGOs. It will strengthen children's and young people's ability to cope with the global societal

impact of technology and provide teachers with pedagogical tools for transforming learning.

Making children and young people aware of the technology on which future societies will increasingly depend is one of the major challenges of the 21st century. Inspired by this challenge, the "Generation AI" project will generate and sharpen the scientific understanding that will enable educational solutions to meet the skills needs of the AI generation.

READ MORE in this issue: Teaching AI skills to primary school students through Generation AI applications

Web-site:

<https://www.generation-ai-stn.fi>

Teaching materials:

<https://www.generation-ai-stn.fi/materialit>

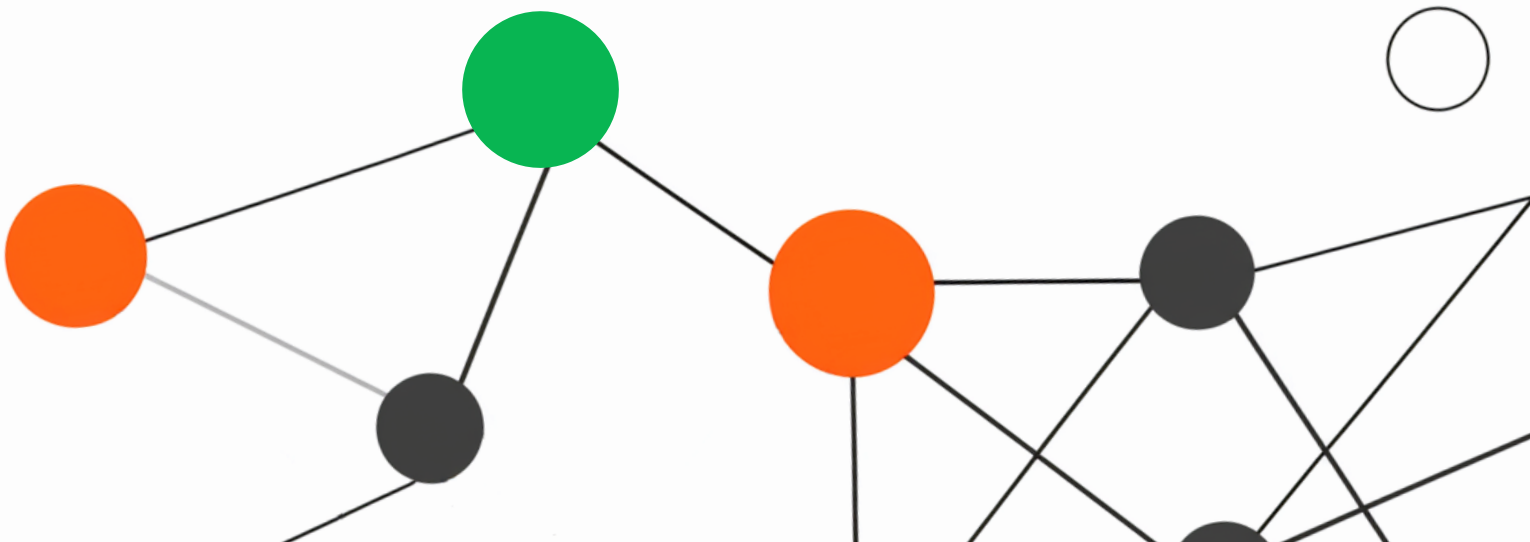
Applications developed for the project:

Teachable machine:

<https://tm.generation-ai-stn.fi>

Social media machine:

<https://somekone.generation-ai-stn.fi>



12. Artificial intelligence changes pedagogy in Lappeenranta

MIIKA MIININ, EDUKS

TWO YEARS AGO, artificial intelligence took the classroom by storm. Students quickly embraced the new tool and teachers were forced to rethink their teaching. The tool, which on the surface looks like a technical gadget, turned out to be a game-changer for the whole school world.

In the autumn, AI passed the matriculation exams with flying colours in several subjects. The answers to the questions were produced simply by entering a question into an AI tool, meaning that no skills were required to answer it with AI. This raises the question of what and how to study in upper secondary schools.

To manage this revolution in which AI would complete assignments for students, the teachers in Lappeenranta sought help. With funding from the Finnish National Agency for Education, Eduks, Otavia, Kuopio High School and Mobie, we responded by producing a guide with 10 concrete tips for assignments in the AI era¹.

The guide does not take a position on whether the right approach is to try to limit the use of AI or to integrate it into the learning process: the guide provides tips for teachers, regardless of their approach to AI.

After the initial shock, teachers began to see AI as an opportunity rather than a bogeyman, and they hoped it could be useful in their own work: how can AI help me do my job better? The next guide continued along the same line with 10 concrete tips. This time, it focused on the background processes that consume teachers' time and energy and reduce their most important work: interacting with students. The tips provided help with communication, brainstorming, producing new ideas, meeting practices, etc.

Accountability only comes into play once the fires have been put out. We are currently working on a guide for teachers on the ethical and safe use of AI. It is a guide that becomes useful now that support for the earlier problems has been provided.

All the above guides, as well as future ones, can be found here (at present only in Finnish):

<https://tekoalyoppimisentukena.wordpress.com/tot-opaat/> (in Finnish)

Staying grounded: AI that respects your data

Many AI services suffer from two basic sins when processing data fed into them: hallucination, i.e. the text contains errors due to an AI guess gone wrong. The second problem is escaping from the data: the AI is trained with a huge amount of text, and the answers stray from the easy-to-process material to the information found in the training material.

ChatGPT, Copilot, Claude and other generative AI tools have revolutionised the way we work in many ways. However, one less noticed AI service has emerged as our favourite in Eduks, especially for tasks that require processing large amounts of data, searching for information or quickly grasping a new thing in a cursory way.

Google's Gemini-based Notebook LM stands out, at least for now, for its ability to stay on top of source material. Notebook LM can be fed with huge amounts of material: documents, web pages, YouTube videos and even raw text. Once you've entered the material, you can chat with the

**Google's
Notebook LM
stands out for
its ability to stay
on top of source
material.**

service, just like any other AI. Notebook LM searches the input materials (and only the input materials) for the desired points, summarises the content, looks for contradictions or overlaps and, if necessary, turns the material into a conversation, where

human-sounding machine voices run through the material like any industry podcast.

Notebook LM offers many interesting practical applications: if an educational institution wants to introduce a new curriculum, then there must be an impact assessment followed by a possible agreement on the processing of personal data. This background work requires a great deal of research, as data protection regulations, legal texts, city guidelines, policies, etc. need to be consulted. The work is speeded up enormously when the necessary materials are fed into the

Notebook LM and, in discussion with it, the necessary passages can be found in the blink of an eye, with clear references.

On the other hand, schools can use Notebook LM for curriculum work, for example, by entering the school's old curriculum, curriculum criteria and the necessary information about the school's priorities. This can help to identify gaps, find out how to effectively integrate the school's priorities into the curriculum and ensure that nothing is forgotten or contradictory.

AI: which would be the better president?

Lecturer Hanna Sydänmaanlakka explored the potential of artificial intelligence in her Academic Writing in English course. In Sydänmaanlakka's example, the assignment required the use of AI: AI was made a tool for processing and internalising factual content.

Students had to feed their chosen AI service ten sources, as different from each other as possible, of the Trump and Harris presidential campaigns. Based on these sources, the students had to form an opinion on which candidate they would prefer, which would be better for Finland and why. In order to justify their choice, the students had to reflect on the reasons behind their choice and compare the candidates in terms of the chosen reasons, policies and promises in discussion with the AI.

Through this hands-on experiment, students identified differences between various AI services: some AI tools refused to discuss politics at all, some strayed from the source material in their responses, and some were just otherwise hallucinating. Students were required to write a paper on their conclusions, name the resources used in the paper, list the AI services used, and analyse how the discussion with the AI went. The time for the whole task was 2 x 75 minutes.

The assignment guided the participants to think

**Schools can use
Notebook LM by
entering the school's
old curriculum.**

about this angle: To choose a better candidate for Finland, one must think about which areas are most important for Finland: economy, security, tolerance or something else? How much does each area weigh in the balance and does one's own opinion differ from that of Finland - why and in what respects?

Critical literacy was developed almost unnoticed during the exercise, as it was easy to find highly polarised views with a strong agenda on the same issue. Sydänmaanlakka was delighted with the quality of the students'

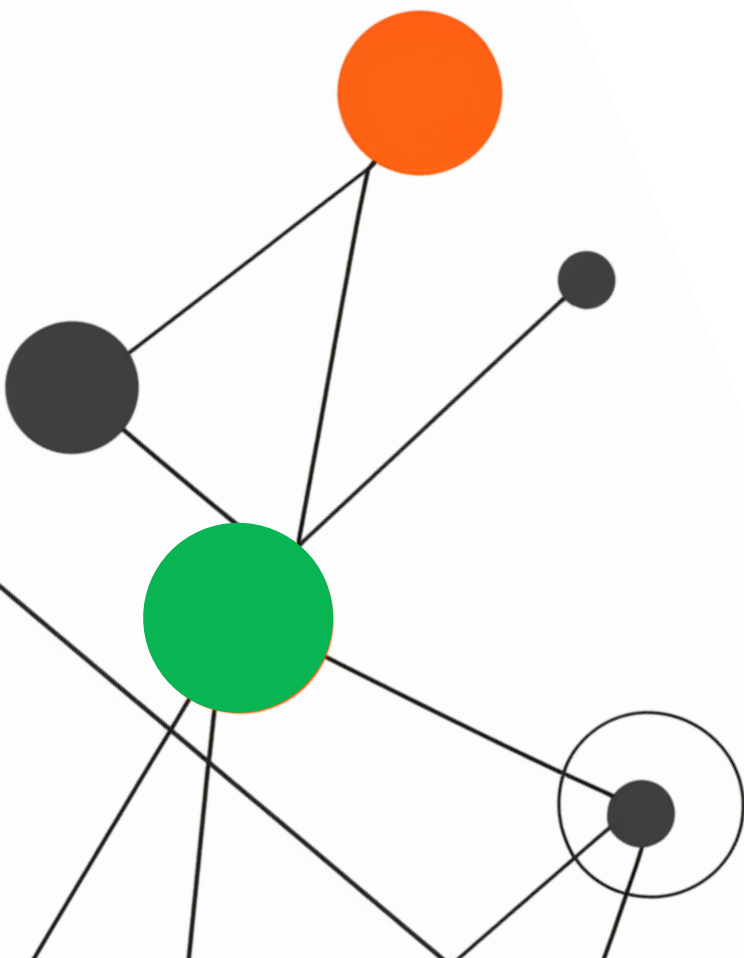
AI is not something to be afraid of.

reflection and reflection. One of the students commented on the assignment thus: "I thought the task would be done in

no time with AI, but it was so interesting that I spent much more time on it than planned!" The experiment also encouraged students to work with AI. Sydänmaanlakka told of a student who had commented that AI had scared him in the past because you could talk to it like a real person. With this assignment, he realised that AI is not something to be afraid of, but that it can be used to analyse data.

AI allows humans to process much larger amounts of source material much faster than traditional methods. When used properly, this leads to better results whatever the task. In other words: more, better, with less.

¹ <https://hankejulkaisut.mobiezine.fi/zine/14/cover>



13. Denialism vs. techno-optimism: a teacher's perspective on AI

AKI SAARIAHO, OTANIEMI UPPER SECONDARY SCHOOL

OUR SOCIETY HAS CHANGED irreversibly – or has it? AI crept into our everyday life and the world of education, partly stealthily, partly with a bang. When talking about AI, we always remember to mention that AI has been part of our devices and applications for a long time¹. On the other hand, AI suddenly came into the midst of education in one way: with ChatGPT in November–December 2022.

For many teachers, generative AI has changed the way they plan lessons, approach learning, and learn with their community of learners. Numerous guides and explanations of AI and its uses are being

There are three main attitudes towards AI applications in learning: denialist, cautious and techno-optimist.

written and produced in various places²³⁴⁵. GenAI tools help with lesson planning and construct content. Pupils and students are happy to take advantage of the possibility of a miracle machine that gives supposedly good-sounding answers. Today – faster than ever before in human history – teachers are learning to use tools at the same time as learners⁶.

Three attitudes

There are three main attitudes to learning about AI applications: denialist, techno-optimist and cautious. These can certainly be broken down in more subtle and specific ways, but these are probably the main lines.

The denialist has decided that there are so many risks, uncertainties and fears associated with AI that they refuse to engage with AI applications altogether. They refuse to use AI in teaching, forbid students from using it in their work, or refuse to learn new applications.

The denialist invokes environmental concerns, highlights the risks of data retention, or worries that humans will forget various skills that we have learned over time if we rely on AI. For the denialist, every litre of computational water consumed by an AI application is proof that the use of AI is unsustainable from the outset, and nothing will make it otherwise.

There is no sustainable way to build data centres or to produce energy without consuming the environment, they say, so it is not worth

trying to improve AI. Literacy and writing are threatened and will die out if we use more AI applications, because like any intoxicating product, they lead to addiction and decay.

The techno-optimist sees AI as an opportunity to take human potential to the next level. AI technology enables a productivity trap by freeing us from trivial tasks towards new spheres of creativity.

The techno-optimist sees only opportunities in AI applications, trusts in risk management and looks forward to the next steps. They do not wait for common government guidelines or national directives but fearlessly explore the possibilities of AI in education and as an individual. For the optimist, this new technology is enabling things and producing previously impossible applications across disciplines and technologies. They cite medicine, transport, personalised learning and rescue from places previously inaccessible as examples.

The cautious one stays firmly within set boundaries. They argue that experimentation can only take place once you have defined what is safe to experiment on. If we don't know where learners' data – any data, such as writing, videos or personal data – is stored, the cautious one will not dare to experiment with AI applications in education. The possibility of personal data ending up in the wrong hands is too great a threat. Permission for AI use should come from the managerial level.

Those with a cautious approach will scan forums and discussion boards before experimenting. They only trust AI after a trusted source has informed them that a particular application is safe to use. Varovainen wants us to have considered all possible agreements and to anticipate risk spots. Cautious is scared by the out-of-control use of AI⁷.

Why use AI and what to watch out for

There are already many different AI applications for different purposes, and more are being

created all the time⁸.

At this point, between 2024 and 2025, we are certainly only at the beginning of the AI applications era. We

expect great benefits, which have been realised for some, but many are still cautious and afraid to even try. Those who remain denialists are fighting the windmills – AI is here to stay, at least in some form.

Above all, AI should be seen as a utility for huge amounts of data; data so huge that it is not worth trying to work on it by human effort or brainpower. Among other things, AI applications help in extracting what is relevant from the source data in a way that is far superior to the time before generative AI. Of course, the skill of creating prompts is essential here: browsing the source material is only as effective and useful as the prompts you use to guide the application to retrieve the desired information.

As for the material, it is of course as “good” and “real” as what is put there. As mentioned elsewhere in this guide, the material used by AI can be corrupted or the source can be very skewed. This can lead to so-called “hallucination”, which is discussed further elsewhere in this guide.

The key to using AI applications is source criticism. GenAI applications are not purely search engines – although even that line starts to blur when integrating search engines with applications that use language models.

The most merciful approach to teaching and learning is to have a learner's source criticism as the basis of everything. Users should be able to assess the following: if the text generated by an AI application is correct; whether it is appropriate in the context; whether it can be improved, if the data retrieved by the AI application can be trusted and whether it answers the question posed. And finally, whether the image/audio/video recording produced by the AI application is appropriate, correct or saves time – also taking into account ethical considerations.

Above all, AI should be seen as a utility to exploit vast amounts of data.

The key to using AI applications is source criticism. GenAI applications are not purely search engines - although even that line starts to blur when integrating search engines with applications that use language models.

The most merciful approach to teaching and also to learning is to have a learner's working source criticism at the basis of everything: being able to tell tentatively whether the text generated by the AI application is correct, appropriate to the context and can be made even better, whether the data retrieved by the AI application can be trusted and whether it answers the question posed in a prompter way, or whether the image/audio/video recording produced by the AI application is appropriate, correct or saves time - also taking into account ethical considerations.

A final note on ethics

At the time of writing, there are both good and bad examples of the ethical side of AI application development. The bad ones include environmental and legal issues. Evidently, developing and working with AI requires a huge amount of energy and natural resources. Not many forests have yet been cleared in Finland to make way for data centres, but they are certainly on the way.

Furthermore, additional energy production capacity will be needed if we are to sustain the development of AI. For example, some of the big AI developers such as Google and Microsoft, just

bought energy for their own needs from nuclear power plants in the US in 2024⁹. New technologies will require new solutions and possibly more consumption. At the same time, we are in an accelerating state of species loss and loss of diversity, which is exactly what we should avoid.

Legislation is behind AI on many fronts: there are many opinions on GDPR, copyright is unresolved¹⁰ and different countries have different positions. Individual privacy and MyData need to be effectively protected.

On the positive side, we can mention both Europe's investment in the AI Act and GDPR's versatility¹¹, along with the huge drive to maintain the ethical side. This protects us against blindly going off the rails in a technological frenzy of non-intelligence¹².

¹ Karmiva löytö TikTok-sovelluksesta: "Poista se viimeistään nyt", read 10.12.2024

² Matleenan blogi: Generatiivinen tekoäly, read 10.12.2024

³ Prompt Library – AI for Education, read 10.12.2024

⁴ <https://tekoalyoppimisentukena.wordpress.com/tot-oppaat/>, read 10.12.2024

⁵ Tekoäly koulutuksessa – lainsäädäntö ja suositukset | Opetushallitus, luettu 10.12.2024

⁶ Change blindness - by Ethan Mollick - One Useful Thing, read 10.12.2024

⁷ AI Reaches Human-Level Reasoning: Should We Be Worried? | by Fix Your Fin, read 10.12.2024.

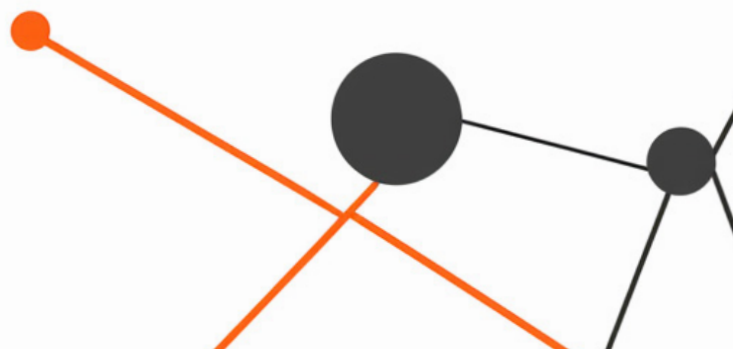
⁸ GenAI opettettava kone → Generation AI, read 10.12.2024

⁹ Why Big Tech is turning to nuclear to power its energy-intensive AI ambitions, read 10.12.2024.

¹⁰ The US Copyright Office says an AI can't copyright its art - The Verge, read 10.12.2024

¹¹ <https://technologymagazine.com/articles/data-privacy-day-data-protection-in-the-age-of-gen-ai%5C>, read 10.12.2024

¹² The impact of the General Data Protection Regulation (GDPR) on artificial intelligence, read 10.12.2024



14. Deepfakes mimic reality

PIPSA HAVULA, FAKTABAARI

WITH GENERATIVE AI, it is possible to very realistically change people's faces, bodies and voices to make them do and say things they have never actually done or said. AI can be used to create images that look like real photographs of anything, or videos that look like real photographs of anything. When this kind of AI technology is used to mimic reality, it is called deepfake.

Deepfakes make use of deep neural networks. They are machine learning models that mimic the principles of biological neural networks. The technology has advanced in leaps and bounds in recent years, thanks in part to the GAN model introduced by US AI researcher Ian Goodfellow and his team in 2014¹. In this model, one neural network acts as an evaluator of another and aims to distinguish genuine outputs from those created by a machine. In this way, the neural networks teach themselves. The goal is that the generating network learns to produce fakes so good that the evaluating network no longer distinguishes them from genuine materials. Increasingly powerful deep forgery applications of the GAN model are still being developed, and forgeries are also being made using other models.

From the entertainment industry to the dissemination of disinformation, deepfakes are used for a wide range of purposes. Deepfake technology can be used to create a digital twin of anyone, for example a deceased actor, who you want to bring back to the big screen. It is also widely used, for example, in the audio book market, where an AI voice reads a book instead of a real person, sounding completely real. And in the advertising market, where the same advertisement can be easily translated into, say, nine different languages.

Deepfake technology has also enabled a lot of bad things to happen. It can be used to create, for example, a sex video that looks like the real thing, without the person's consent, from any person about whom images or video material is available. It is now widely used to shame people, especially women and girls, all over the world. Doctored, fake-looking videos are also used to bully schools and put pressure on journalists².

Human voice cloning has also been used to cheat parents out of money by calling them with their own child's voice³. And tricked

a company employee into transferring a large sum of money to an account where it should never have been transferred in the first place⁴. Examples abound.

Today, almost without exception, deeply distorted images, videos and audio tapes are also distributed during major news events. In elections, deep fakes have been used to blind politicians and undermine confidence in electoral systems. In the US, an AI voice cloned from Joe Biden called people in the run-up to the 2024 presidential election and urged them not to vote⁵. In the same election, an attempt was made to paint Tim Walz as a paedophile by publishing a deeply distorted video online in which a fake “former student” of Walz spoke to a camera about his horrific experiences⁶. The list of such examples is endless, and Finland has not been immune to fakes.

In order for deepfakes to cause damage, they need to be able to be spread via messaging services and social media. Digital giants such as Meta, Tiktok, Google and Microsoft have been called upon to take responsibility for curbing the spread of deep counterfeiting.

For deepfakes to cause damage, they must be able to spread through messaging services and social media. Digital giants such as Meta, Tiktok, Google and Microsoft have been called upon to take responsibility for curbing the spread of deepfakes.

Identification of deep fakes

Today, it is almost impossible to distinguish a cleverly made deepfake from the real thing. The same can be said in reverse: it has become increasingly difficult to authenticate genuine video as genuine video. As a result, in many situations, perfectly authentic material has been accused of being faked. This can have serious consequences for the dissemination of correct information and for what people ultimately believe.

Until recently, you could identify a deepfake if an AI created too many or too few fingers for people, or if it stumbled over earplugs and

glasses. Gradually, however, such small errors have been ironed out, and deepfakes created with good AI applications are now hard to distinguish from the real thing. Voice forgeries have also become almost flawless.

The general guidelines below will help you assess the reliability of information you receive online - even when it is unclear whether or not the material is deeply distorted:

- **If an image, video or audio clip you receive on social media evokes a strong emotional response, the first thing to do is to stop, think and think again before believing the material and sharing it.**
- **Consider three questions. Who has posted the photo/video/audio tape? What evidence has been presented to support it? What other sources have to say about it?**
- **Use a lateral reading approach: have reliable sources reported on this? Has anyone questioned the authenticity of the material in the comments? Can you find other photos/videos/material from the same event?**
- **Put a screenshot of the image or video into a reverse image search (e.g. Google Lens) and see where the footage has been circulated in the past.**
- **Be sceptical, especially if a scandalous photo, video or audio tape starts circulating on social media just before the election. If necessary, contact the authorities, fact-checkers or other journalistic media.**

A solution to this situation has also been sought in the form of detection technologies, which can be found online under the search term “deepfake detector”. However, the FaktaBaari does not recommend relying on them alone. There is a constant race between the development of deepfakes and their detection, with detection technologies always one step behind⁷.

Identification tools may mistakenly mistake genuine material for counterfeit, or counterfeit material for genuine⁸. Algorithms that try to detect fakes are

only as good or bad as the data they are trained on. Recognition technology also has its biases and may not recognise fakes of people with dark skin, for example, or be confused if a genuine image or video has been shrunk or even blurred⁹. The quality of recognition tools varies enormously, and just as money is made with all other AI tools, it

is also made with recognition tools.

Thus, identification tools can be tried as part of the fact-checking process, but they should be treated as helpers rather than experts.

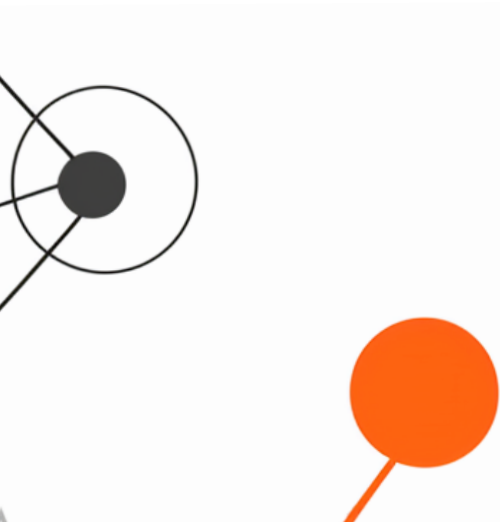
It is worth trying to

familiarise yourself with the source material on which the identification tool has been trained, who developed it and what its limitations are. Never take the assessment of the detection tool as fact, but also look for information from other sources.

As with any other technological development, it is worth keeping your knowledge up to date with identification tools.

There is a constant race between the development of deepfakes and their detection. The detection technology is always one step behind.

- 1 Cf. MIT Technology Review (2018). The GANfather: The man who's given machines the gift of imagination <https://www.technologyreview.com/2018/02/21/145289/the-ganfather-the-man-whos-given-machines-the-gift-of-imagination/>
- 2 BBC (2024). Inside the deepfake porn crisis engulfing Korean schools <https://www.bbc.com/news/articles/cpdlpj9zn9go>; BBC (2024). Girl, 12, victim of 'deepfake' bullying porn image <https://www.bbc.com/news/articles/ckvgez74kgo>; VOA (2024). Journalist 'haunted' by AI deepfake porn video <https://www.voanews.com/a/journalist-haunted-by-ai-deepfake-porn-video/7624281.html>
- 3 Guardian (2023). US mother gets call from 'kidnapped daughter' - but it's really an AI scam <https://www.theguardian.com/us-news/2023/jun/14/ai-kidnapping-scam-senate-hearing-jennifer-destefano>
- 4 Guardian: Company worker in Hong Kong pays out £20m in deepfake video call scam <https://www.theguardian.com/world/2024/feb/05/hong-kong-company-deepfake-video-conference-call-scam>
- 5 Reuters (2024). Consultant fined \$6 million for using AI to fake Biden's voice in robocalls <https://www.reuters.com/world/us/fcc-finalizes-6-million-fine-over-ai-generated-biden-robocalls-2024-09-26/>
- 6 Wired (2024). Russian Propaganda Unit Appears to Be Behind Spread of False Tim Walz Sexual Abuse Claims <https://www.wired.com/story/russian-propaganda-unit-storm-1516-false-tim-walz-sexual-abuse-claims/>
- 7 Binh, Le, Shahroz Tariq, Sharif Abuadbba et al. (2023) Why Do Deepfake Detectors Fail? https://www.researchgate.net/publication/368843207_Why_Do_Deepfake_Detectors_Fail
- 8 Cai, Zhixi and Ghosh, Shreya and Adatia, Aman Pankaj et al. (2024). AV-DeepfakeIM: A Large-Scale LLM-Driven Audio-Visual Deepfake Dataset <https://arxiv.org/pdf/2311.15308>
- 9 University at Buffalo. (2024) New deepfake detector designed to be less biased <https://www.buffalo.edu/home/story-repository/host.html/content/shared/university/news/ub-reporter-articles/stories/2024/01/tyu-deepfake-bias.detail.html>



15. AI-based recommendation algorithms for social media services

HARTO PÖNKÄ, INNOWISE

WHEN PEOPLE TALK ABOUT AI, the first things that come to mind are chatbots like ChatGPT and Copilot, and image generators using AI models. At the same time, it may go unnoticed that we use AI much more often in non-AI applications: namely in social media services.

In the most used social media services, machine learning has been used to make algorithm recommendations since 2016 at the latest. The earliest experiments with recommendation systems were done years earlier.

Different recommendation algorithms

Recommender systems are used in a wide range of information systems and applications. In social media services, they are commonly referred to as recommendation algorithms or personalised algorithms. These are programmes that use machine learning to sift through large amounts of data to generate recommendations that are appropriate for users. Most commonly, it is the generation of content feeds that are visible to users on a social networking site. In effect, recommendation algorithms decide what content to show users on social media services and applications and in what order.

Recommendation systems can be mainly divided into three categories: content-based, collaborative and hybrid models that combine several approaches.^{1 2}

- **A content-based recommendation algorithm makes recommendations by comparing information provided to users and content. For example, the app can ask users what topics they are interested in and then display content that matches the selected categories.**
- **A collaborative recommendation algorithm makes recommendations based on data collected from users' activities. The idea is that instead of a single user filtering content, data from multiple users is used to filter content, and users who have been interested in the same topics in the past will be interested in the same topics in the future. The data can include, for**

example, user reactions to posts and the user's social network.

- **Hybrid recommendation algorithms combine content-based and collaborative approaches with other methods. For example, an algorithm can recommend content to new users based on their location or interests and then move to collaborative recommendations as data about the user's activity accumulates and can be compared with data about other users and content.**

As users' interests, behaviour and activity levels vary, there has been a growing effort to develop algorithms that are more adaptable to their habits and differences. This is reflected in the proliferation of hybrid recommendation algorithms. For example, a very active social media user will need a much larger number and more recent recommendations than a user who visits an application, say, only a few times a week.

Existing online and social services aim to collect the most accurate information possible about users' behaviour, so that recommendation algorithms can make the best possible choices for users. This has meant that more and more personal and behavioural data is being collected from users.

As the amount and complexity of user data has increased, applications have developed AI-based recommendation algorithms. The advantage of AI-based algorithms is that they can make efficient use of a large pool of user-specific data and adapt to each user's behaviour. This means that different variables can be weighted when making recommendations, depending on the data

available about the user.

This is a clear advantage over rigid recommendation algorithms that follow the same rules for each user.

Existing recommendation algorithms use several AI techniques such as machine learning, deep learning, and natural language processing

familiar from conversational bots. On the other hand, despite the use of AI, users may feel that the choices made by social networking algorithms are not suitable for them.

It is good to be aware that the algorithms of online and social services change frequently. These changes are the result of algorithms evolving as social media services seek to optimise their performance to suit their purposes. For example, as artificial intelligence develops, there is a growing effort to use new AI functions in recommendation algorithms.

Whether or not AI is used, it is worth bearing in mind two things that seem to be relatively constant: first, algorithms are generally designed to keep users on the service as long as possible, and second, algorithms are susceptible to manipulation. For example, advertisers, political influencers and disinformation disseminators are constantly studying algorithms to maximise the visibility of their publications.³

For better or worse, Facebook has been a pioneer of AI algorithms. Facebook launched its News Feed two years after the launch of the service in 2006. Initially, there was no recommendation algorithm, but the News Feed contained all the activity of Facebook friends in chronological order.

Facebook released its first algorithm to drive the News Feed in 2009. In retrospect, this event can be seen as the beginning of the social networking recommendation algorithms. The idea was to show users the most interesting content instead of showing them all the posts of their friends. The algorithm mainly used three variables: the proximity of the user to the post, the weight of the content and the time of the post. It was therefore a collaborative recommendation algorithm where the social network between users played a big role. Later, this type of algorithm came to be regarded as a standard one, specifically for networking services.^{4 5 6}

The introduction of the algorithm was a significant change for Facebook's 300 million users at the time. By no longer "having" to read

An active social media user needs more recent recommendations than a user who only visits the app a few times a week.

all their friends' posts, users were able to network with more and more people. In 2009, the average number of Facebook friends was 120, but by 2011 the number had grown to 190. That's when another major change was made to the News Feed algorithm, with the introduction of a machine learning-based recommendation algorithm⁷.

By 2013, the new AI-based recommendation algorithm used up to 100,000 different variables. This meant increasingly personalised recommendations: users who were active in Facebook groups saw more group posts. Those who frequently liked photos saw more and more photo posts. In turn, those who clicked on links were offered more links, and so on⁸.

As AI sifted through the most interesting posts on the social network, users were able to network with more and more of them without cluttering up the News Feed with uninteresting "noise". In 2016, the average US user already had 350 Facebook friends, and as many as 650 for 18-24-year-olds.

In the early years, Facebook had been a platform for communication between friends, but the growth in friends meant the possibility of ever greater publicity. The importance of links shared on Facebook became even more important for commercial entities, as organic visibility on Facebook could mean thousands and thousands of visitors to a company's website. Alongside social networking, the motives for using Facebook may have become mass appeal, business and political influence. Facebook's own employees described the change as a collapse of context⁹.

In 2016, Facebook released emoji reactions of "love", "haha", "wow", "sad" and "angry". These allowed the algorithm to infer users' emotional state and recommend posts that matched it. At the same time, the hybrid AI recommendation algorithm became so complex that it proved difficult for Facebook to manage.

There was one serious problem with the effect of emoji reactions on the recommendation algorithm. Emoji reactions were given many times the weight of normal likes when calculating the algorithm's recommendation score. As a result,

posts with strong reactions such as "angry" and "haha" emojis could gain disproportionate visibility when the algorithm picked them up in user feeds. Thus, triggering anger and other strong emotions became an effective way to get posts wide visibility on Facebook. This did not go unnoticed by those seeking^{10 11 12}.

It took Facebook years to realise and correct the disproportionate impact of emoji reactions on the functioning of the recommendation algorithm. In 2021, according to documents leaked by a former Facebook employee, Facebook had tried to prevent emotional reactions from over-influencing recommendations, but a feature developed for this purpose did not work. Tellingly, the News Feed algorithm could give high visibility to posts containing disinformation, hate speech and clickbait, for example, while the company's moderators did their best to weed out such content.

Facebook researchers discovered the use of emoji reactions to spread inaccurate and harmful content in 2019. The impact of angry reactions on the recommendation algorithm was calculated several times and finally reset in 2021 after Facebook came under heavy public criticism in the aftermath of the 2016 and 2020 US presidential elections. The company was accused of failing to combat disinformation and election interference. Emotional Facebook posts were also used effectively in the campaign for the 2019 Finnish parliamentary elections¹³.

After 2021, Facebook has continued to develop its AI-driven recommendation algorithm. With increasing data protection regulation, Facebook has started to regularly publish information on which posts the algorithm selects for the average Facebook user's feed. In recent years, Facebook has reduced the visibility of links in particular, and increased the number of posts whose publishers are not part of the user's network of Facebook

Generating anger and other strong emotions became an effective way of getting posts widely seen on Facebook.

friends. These changes mean that Facebook is increasingly seeking to keep users on its service and is gradually moving towards interest-based recommendation based on profiling, where users' social network plays an increasingly less important role¹⁴.

YouTube's problem: which of the millions of videos to recommend next?

YouTube's need for an efficient recommendation algorithm was obvious when, just three years after the service's launch in 2008, more than 45 million videos had been uploaded to the site. Videos could be browsed by channel, ranked by popularity, and via text-based search. However, the search function did not work very well, as for most videos the only search criteria that could be used was the title of the video¹⁵.

The closest to YouTube's need to develop a recommendation algorithm was a competition sponsored by Netflix in 2006–2009 to develop the most accurate movie recommendation system possible. Netflix had started renting DVD movies by mail and had a database of around 100,000 DVDs at the time. The company promised \$1 million to anyone who could develop a recommendation system that was at least 10% better than its algorithm at the time. Compared to Netflix, YouTube had a much larger selection of videos and less information about its users.

Google started developing machine learning for video recommendations on YouTube in 2007. The company's researchers decided to create a network of videos based on how many times the same users had watched the same video. Presumably, it made sense for users to recommend videos that had been watched before by users interested in the same videos. The challenge was the large number of videos, which made it impossible to calculate new video recommendations for each user every time they switched between different videos. On the other hand, the aim was not to suggest only the most popular videos to users, but to allow them to find new content of interest.

As a solution, YouTube's developers created several AI models that used log data from users as machine learning data. The AI models were able to suggest various "related videos" to users based on videos they had just watched. This was the beginning of one of the early collaborative filtering-based social media recommendation algorithms. Unlike Facebook, the recommendation was not based on the social network of users, but on a network of videos.

YouTube has been a pioneer in the development of AI functions beyond the recommendation algorithms. In 2009, YouTube launched automatic video subtitles based on speech recognition, and in 2013, facial recognition for people and animals, such as cats.

It is unclear how long YouTube's early AI algorithm was used to recommend videos. However, in 2016, a more sophisticated recommendation algorithm using deep learning and neural networks was introduced. The new AI-based algorithm allowed for more efficient moderation of content to automatically identify content that was e.g. inappropriate or otherwise harmful to children. Since 2017, YouTube has been working to boost content from trusted media in the algorithm. In 2019, YouTube further refined the algorithm to reduce the number of recommendations for videos containing questionable claims.^{16 17}

YouTube has been a pioneer in the development of artificial intelligence functions.

TikTok's addictive algorithm

TikTok is new compared to many other social networking services. Its sister app Douyin was launched in China in 2016, and TikTok in other countries the following year. In Finland, TikTok became particularly popular between 2019 and 2021¹⁸.

TikTok's parent company ByteDance is primarily an AI company. One of its first products in 2012 was the Chinese news recommendation app

TikTok's "invention" was not short videos as a form of content, but a recommendation algorithm that managed to capture the interest of users.

Toutiao, based on a user-personalised recommendation algorithm using machine learning and natural language processing. In 2015, short videos were added to the app, and the following year Toutiao was already the most popular short video service in China¹⁹.

Unlike many other social media services, ByteDance had a powerful recommendation algorithm in place when it launched Douyin and TikTok. It is worth noting that TikTok's "invention" was not short videos as a form of content, but rather a recommendation algorithm that managed to capture the interest of users. At the time, there were already many widely used short video apps such as Vine (2012-²⁰ ²¹). TikTok combined public short videos with an AI algorithm based on profiling, which recommends videos of interest to users based on data.

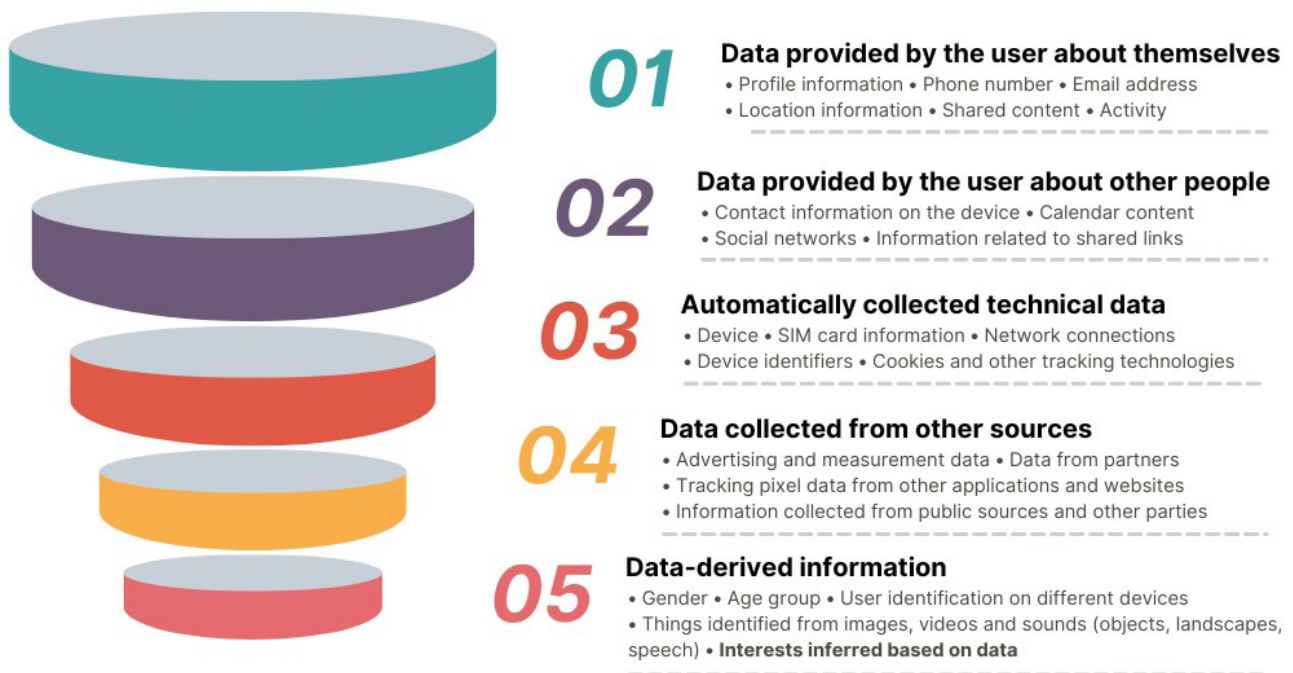
The core of TikTok is the For You Page, which is made up of recommendations from the algorithm.

The typical use of the application is to quickly browse the feed and watch interesting videos. Since viewing decisions are based on the video's opening image, title and start, visual interest and attention-grabbing are key. This is underlined by the fact that, like many other social media services, TikTok's algorithm is not based on which publishers the user has tracked, but on what topics and what kind of videos the user has watched in the past.

According to TikTok, the following factors, among others, influence the recommendation of videos in the For You feed:

- **User interactions: videos liked, shared, commented, watched, skipped and accounts followed.**
- **Video data: voices, hashtags, captions, view counts and countries of publication.**
- **User data: device settings, language, location, time zone and day, and device type.**

Compared to other social media services, TikTok's data collection is very extensive. TikTok's data collection can be divided into five levels as shown in the figure below²²:



Harto Pönkä, 28.1.2025

Levels of data collection in TikTok

1. User self-reported data
2. Data provided by the user about other people
3. Automatically collected technical data
4. Data collected from other sources
5. Data derived from the data

By comparing the data collected from the user and the video data, the likely interest of each video to the user is determined. Different factors are weighted in different ways. For example, watching a video from start to finish is a strong indicator that a user is interested in the content of a particular topic. The mere fact that a publisher has many followers does not make their videos superior to others in the recommendation system. The social network, i.e. the number of followers, does not play a major role in the operation of the TikTok algorithm. The TikTok algorithm can be characterised as a recommendation system driven by the interest network.

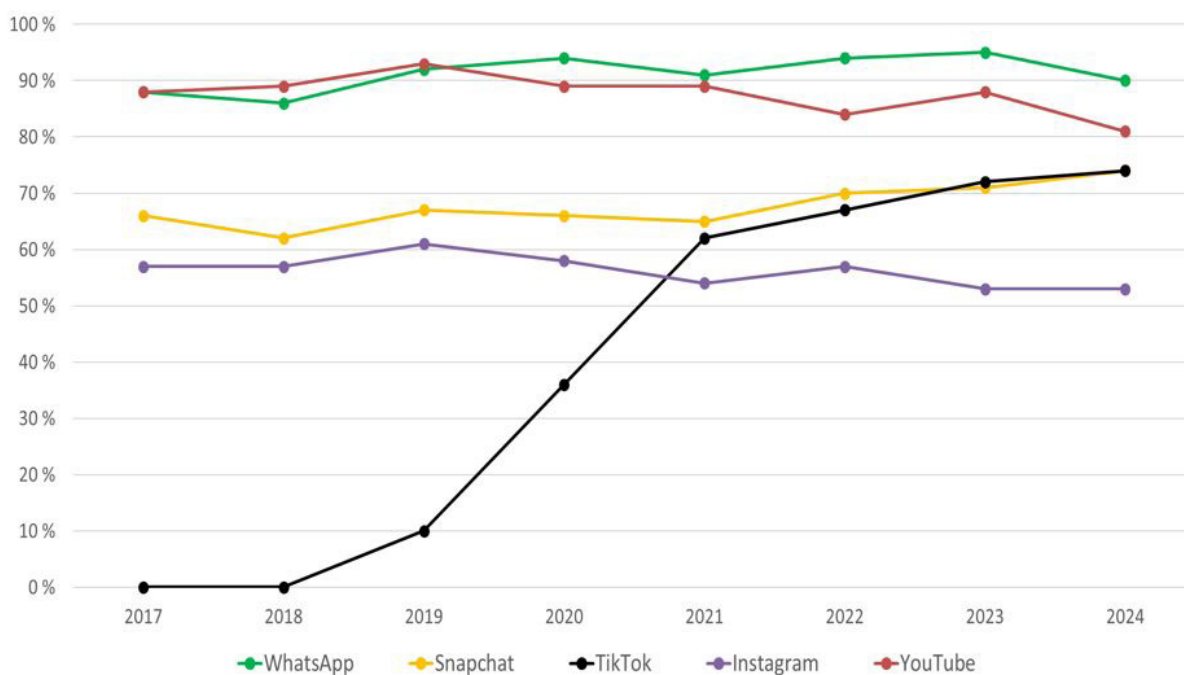
TikTok's algorithm is very reactive: the user's actions on the platform quickly influence the videos that are shown to him in the future. This makes TikTok very addictive for its users. According to

internal company documents leaked in 2024 about the lawsuits filed against TikTok in the US, users could be hooked in 35 minutes and fall into a filter bubble created by the algorithm in as little as 20 minutes²³.

Short videos are particularly appealing to young people

The worldwide popularity of TikTok has set the pace for the development of social media services in recent years. As with many other social media services, the growth in TikTok's user base was initially driven mainly by children and young people, who quickly adopted the new application.

Before the rise of TikTok, the most popular social media services among young people in Finland were WhatsApp, YouTube, Snapchat and Instagram. The popularity of WhatsApp and Snapchat has been based mainly on instant messages and groups between users, video content on YouTube, and photo content on Instagram. TikTok's short videos offered a new type of content compared to previous social media services and spread rapidly among young Finns from 2019 onwards. By 2021, TikTok was already more popular than Instagram and on par with Snapchat among 13–16-year-olds.



Datalähde: DNA, Koululaistutkimus, 2017-2024, käyttö päivittäin, lähes päivittäin tai useita kertoja viikossa, kuva: Harto Pönkä, 30.12.2024.

Most popular social media services for 13–16-year-olds in Finland 2017–2024

However, TikTok's entry into the market does not seem to have reduced the use of other popular social media services by young people much.²⁴

TikTok's popularity has not gone unnoticed by other social media giants. With the release of Instagram's Reels videos in August 2020 and YouTube's Shorts videos in July 2021, TikTok can be considered the biggest challenger to Instagram and YouTube as a video service – especially among young users. It is therefore understandable that they are trying to offer short videos similar to TikTok.

Short video services have changed and continue to change the way people use social media. This is reflected, for example, in the fact that in just a few years short videos have become more popular in many countries as a way of following the news than long videos or live video broadcasts. According to a study by the Reuters Institute (2024), news videos are most viewed globally on YouTube, Facebook, Instagram and TikTok. Instagram and TikTok are particularly popular among people under 35. TikTok is also becoming an increasingly important news channel in the Global South.²⁵

The question of news monitoring is whether there is a difference in the reliability of news received through different channels. According to several surveys, TikTok users often perceive the news they receive via TikTok as unreliable. In a Reuters survey, the most frequent difficulty in verifying the reliability of news on TikTok was perceived to be the difficulty of verifying the reliability of news on TikTok. It was followed by X, Facebook, and Instagram. In Finland, the results matched the average for the countries surveyed.²⁶

In Finland, the effects of TikTok and the short video boom have been felt quickly. According to a survey published in early 2024, TikToki had become the most popular news channel among young Finns aged 13–18, with 33% saying they got all the news they needed. The second most popular news channel was the websites of tabloid newspapers and the third was Instagram. Most²⁷. This is a rapid change, which means that young people's

media consumption is very different from that of most of the adult population. There is concern that a relatively high proportion of TikTok users (35%) could not say whether they had encountered²⁸. This reflects both the need for digital information literacy among young social media users and the difficulty of detecting disinformation on a short video service in particular.

The growing power of data-driven algorithms

While the most visible impact of TikTok is the proliferation of short videos, it's worth remembering that the secret to its success is not short videos per se, but a powerful recommendation algorithm. It is also about the extent to which the content recommendations of the algorithm are used by the service.

The success of TikTok – and YouTube before it – has shown that an AI recommendation algorithm based on profiling and interest networks can be more effective in capturing users' interest than an algorithm based on social networking and collaborative filtering. In TikTok's algorithm, social relationships between users play little role. This is a significant difference, especially compared to the algorithms of Facebook, Instagram and X (formerly Twitter), which were originally built on the social network of users.

Meta has been open about the development of its recommendation algorithms. In response to TikTok's competition, Meta has invested in the use of AI in recent years, for example in its Instagram Reels algorithm. In early 2021, Meta announced that it had developed a continuous self-learning AI model for Reels based on the analysis of video audio and video images.

The results of the image and speech recognition of videos are combined with the analysis of their titles and captions, enabling the AI model to generate more accurate video recommendations

Short video services have changed and continue to change the way people use social media.

that match the user's interests. Like TikTok, Instagram also collects a significant amount of profiling data, i.e. signals related to a user's activity, to infer their preferences.^{29 30}

AI-powered hyper-personalised recommendation algorithms have increasing power over what we see on different digital platforms.

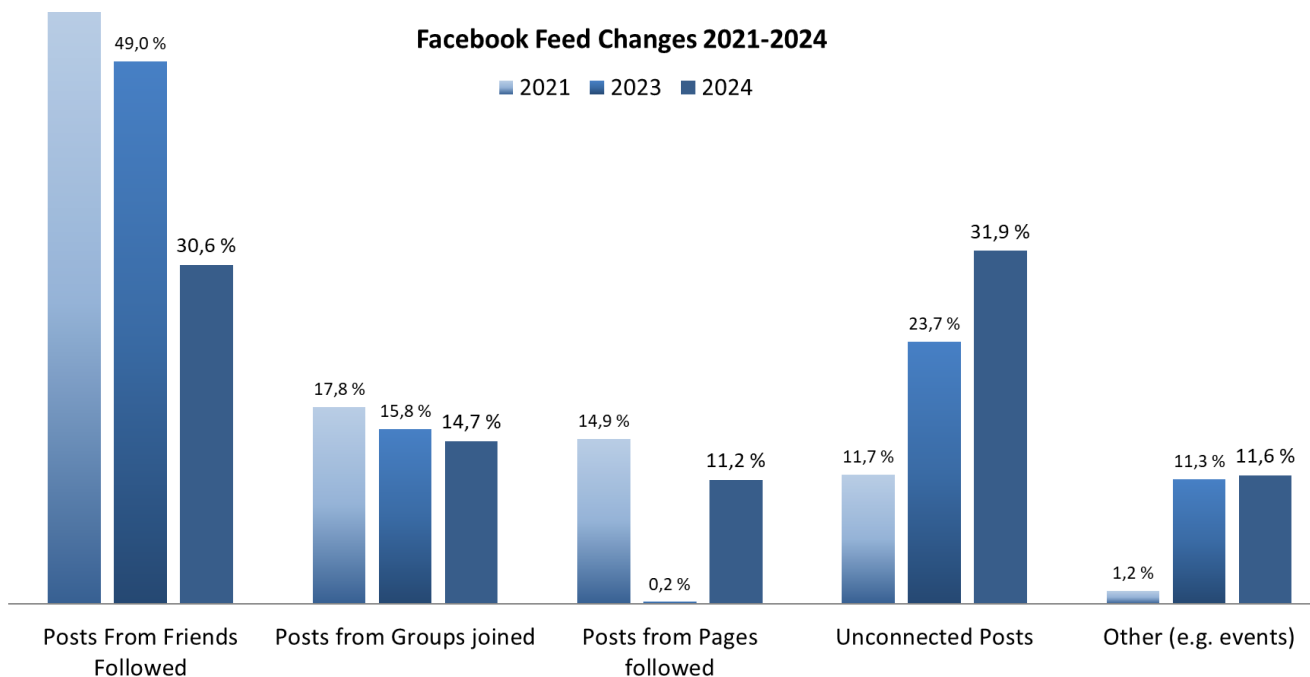
Meta has also increased the use of profiling data in Facebook's feed algorithm. This is reflected in the fact that more and more of the posts selected by the algorithm for the feed are so-called unlinked posts, i.e. posts by publishers that users do not follow. In 2021, only around 12% of the posts in the feed were, on average, from untracked publishers, but in 2024 this figure had risen to around 32%. At the same time, posts from Facebook friends in particular show up less and less in the feed. In other words, users' preferences and social relationships have less and less influence on what they see on Facebook.³¹

An interesting question is to what extent the algorithms of social services respond to users' needs. One answer to this question is provided

by a study by Taylor and Choi, who investigated how the content of different social networking services responded to users' personal preferences. Participants rated the performance of algorithms on Facebook, Instagram, Twitter and TikTok based on a total of 15 statements.

Eight of the statements were related to the desired behaviour of the algorithm and seven to the undesired behaviour. Based on the results, the content selected by the TikTok algorithm was found to reinforce and meet participants' goals more often than Facebook, Instagram and Twitter. The Facebook algorithm was perceived to provide the most recommendations that did not meet users' goals. The study was conducted between 2021 and 2022, before Facebook increased the proportion of non-tracked posts in the algorithm's recommendations³². On the other hand, it is worth noting that participating in research based on algorithms requires a quite advanced level of algorithm awareness and algorithm literacy, so that participants can compare the performance of algorithms in different applications.

The trend in the development of recommender algorithms in recent years can be summarised by



Data source: Meta, Transparency Center, Widely Viewed Content Report 2021 Q3, 2023 Q3 and 2024 Q3, US users, <https://transparency.fb.com/fi-fi/data/widely-viewed-content-report/> Image: Harto Pönkä, 28.1.2025.

Facebook feed changes 2021-2024

the increasing importance of data-driven algorithms compared to socially-driven algorithms. This can be expected to further change the way social media is used in the near future. At the same time, data-driven AI-based algorithms will increasingly influence the behaviour of billions of social media users at both individual and community level³³.

AI-powered hyper-personalised recommendation algorithms mean that algorithms, rather than users, will have increasing control over what we see across social media services and other digital platforms. After all, algorithms are developed according to the goals of platform owners, not users. On the other hand, the above-mentioned research comparing algorithms across different social services suggests that data-driven algorithms may be more responsive to users’ needs than socially-driven ones.

Algorithm regulation bites badly

The impact of digital platforms on users is not limited to the user experience per se. For example, the impact on privacy of data collection by platforms, the veracity of content recommended by algorithms, the ability to block harmful content and hate speech, and ultimately the impact on the overall well-being of users, must be assessed. Many EU regulations such as the General Data Protection Regulation (GDPR), the Digital Services Regulation (DSA) and the Artificial Intelligence Regulation (AI Act) are also relevant. Since 2018, GDPR has had an impact on how digital giants collect and use user data. For example, Meta and TikTok have been fined

hundreds of millions of euros under the GDPR for failures in the processing of children’s personal data.³⁴

The Digital Services Regulation came into force in February 2024, bringing several new rights for users of digital services. The Regulation prohibits the targeting of advertising based on specific personal data such as health, political opinions, ethnic origin or beliefs. In addition, minors may not be targeted based on profiling³⁵.

The Digital Services Regulation requires online services to be transparent about how their referral systems work. A wealth of information on the functioning of the algorithms of different social networking services is therefore available. At the same time, it should be noted that this information is often incomplete and arbitrary. For example, Meta publishes a lot of information on how its algorithms work, but there is little information on the weighting of feeds and the actual amount of content on Instagram and Threads, for example. For those interested in how algorithms work, the most illuminating sources of information remain data leaks, studies and practical experiments.

The Digital Services Regulation introduced an obligation for large online platforms to offer an alternative to the recommendation system that is not based on user profiling. The table below summarises the normal functioning of Facebook, Instagram, YouTube, TikTok and X algorithms and the alternatives they offer to a profiling-based algorithm.

For example, YouTube is still reasonably easy to use, with the ability to track people of interest. The

	Facebook	Instagram	Youtube	Tiktok	X
Publications selected by the algorithm from the user's tracked tags	About 68 %	60-70 %	Minor	Minor	50 %
Algorithm recommendations from publishers not followed by the user	About 32 %	30-40 %	Up to 100 %	Up to 100 %	50 %
Machine learning in algorithm recommendations	2011	2016	2016	2016	2016
Alternative feed	Publications of the followed accounts in chronological order	Publications of the followed accounts in chronological order	Publications of subscribed channels in chronological order	Popular publications by user region	Publications of the followed accounts in chronological order

same goes for Facebook, Instagram and X: all of them allow you to see what the social network has posted recently. In contrast, TikTok's alternative to the AI-driven "You" feed is odd: the user is then shown videos that are popular in their area³⁶. In practice, TikTok's alternative is a random feed that hardly interests anyone. Thus, it meets the DSA requirement, but the benefit to users approaches zero.

TikTok is also a cautionary tale of an application that publicly claims to meet the requirements of the DSA, but whose practical measures do not provide users with the benefits that the DSA is intended to deliver. TikTok has had, among other things, major shortcomings in removing illegal content and blocking underage users. In addition, the transparency requirements of the Digital Services Regulation are of little use if the truthfulness of the reports produced by the company cannot be verified³⁷.

- 1 Wikipedia. (28.12.2024). Recommender system, https://en.wikipedia.org/w/index.php?title=Recommender_system&oldid=1265732754
- 2 Wikipedia. (12.12.2024). Collaborative filtering, https://en.wikipedia.org/w/index.php?title=Collaborative_filtering&oldid=1262596993
- 3 TikTok. (19.6.2020). How TikTok recommends videos #ForYou, <https://newsroom.tiktok.com/en-gb/how-tiktok-recommends-videos-for-you-eu>
- 4 McGee, M. (16.8.2013). EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors, <https://martech.org/edgerank-is-dead-facebook-news-feed-algorithm-now-has-close-to-100k-weight-factors/>
- 5 Facebook. (27.4.2009). Facebook press room, available at: <https://web.archive.org/web/20090427160416/http://www.facebook.com/press/info.php?statistics>
- 6 Wallaroo. (2024). Facebook News Feed Algorithm History, read 9.12.2024, <https://wallaroomedia.com/facebook-newsfeed-algorithm-history/#one>
- 7 McGee, M. (16.8.2013). EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors, <https://martech.org/edgerank-is-dead-facebook-news-feed-algorithm-now-has-close-to-100k-weight-factors/>
- 8 McGee, M. (16.8.2013). EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors, <https://martech.org/edgerank-is-dead-facebook-news-feed-algorithm-now-has-close-to-100k-weight-factors/>
- 9 Pönkä, H. (31.8.2016). Facebook meni rikki, <https://harto.wordpress.com/2016/08/31/facebook-meni-rikki/>
- 10 Pönkä, H. (31.10.2021). Infografiikka: Facebookin viha-reaktio ja algoritmin muutokset, <https://harto.wordpress.com/2021/10/31/infografiikka-facebookin-viha-reaktio-ja-algoritmin-muutokset/>
- 11 Wikipedia. (28.12.2024). Recommender system, https://en.wikipedia.org/w/index.php?title=Recommender_system&oldid=1265732754
- 12 Owen, L. (15.4.2019). One year in, Facebook's big algorithm change has spurred an angry, Fox News-dominated — and very engaged! — News Feed, <https://www.niemanlab.org/2019/03/one-year-in-facebooks-big-algorithm-change-has-spurred-an-angry-fox-news-dominated-and-very-engaged-news-feed/>
- 13 Knuutila, A. & Laaksonen, S-M. (2020). Viraali vihasuus ja tahmea nauru: tunteet ja algoritmit digitaalisessa vaalikampanjoinnissa, julkaisussa S Borg , E Kestilä-Kekkonen & H Wass (toim) , Poliittikan ilmastonmuutos : Eduskuntavaalitutkimus 2019 . Oikeusministeriön julkaisuja. Selvityksiä ja ohjeita , Nro 2020-5, Oikeusministeriö, <https://julkaisut.valtioneuvosto.fi/handle/10024/162429>

- 14 Meta. (2024). Facebook Transparency Center, Widely viewed content report: What People See on Facebook, Q3 2024, read 11.12.2024, <https://transparency.meta.com/fi-fi/data/widely-viewed-content-report/>
- 15 Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S., Ravichandran, D. & Aly, M. (2008). Video suggestion and discovery for youtube: taking random walks through the view graph, <https://static.googleusercontent.com/media/research.google.com/fi//pubs/archive/34407.pdf>
- 16 Youtube. (15.9.2021). On YouTube's recommendation system, <https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/>
- 17 YouTube. (n.d.). Videosuosittimet: Miten YouTube suositusjärjestelmä toimii?, read 11.12.2024, https://www.youtube.com/intl/ALL_fi/howyoutubeworks/product-features/recommendations/#signals-used-to-recommend-content
- 18 Wikipedia. (31.12.2024). TikTok, <https://en.wikipedia.org/w/index.php?title=TikTok&oldid=1266416079>
- 19 Wikipedia. (7.9.2024). Toutiao, <https://en.wikipedia.org/w/index.php?title=Toutiao&oldid=1244468033>
- 20 Wikipedia. (26.12.2024). Snapchat, <https://en.wikipedia.org/w/index.php?title=Snapchat&oldid=1265349499>
- 21 Wikipedia. (21.11.2024). Vine (service), [https://en.wikipedia.org/w/index.php?title=Vine_\(service\)&oldid=1258710466](https://en.wikipedia.org/w/index.php?title=Vine_(service)&oldid=1258710466)
- 22 Pönkä, H. (29.4.2024). Uusi selvitys TikTokin riskeistä ja raportti TikTokin Kiina-yhteyksistä, <https://harto.wordpress.com/2024/04/29/uusi-selvitys-tiktokin-riskeista-ja-raportti-tiktokin-kiina-yhteyksista/>
- 23 NPR. (11.10.2024). TikTok executives know about app's effect on teens, lawsuit documents allege, <https://www.npr.org/2024/10/11/g-si-27676/tiktok-redacted-documents-in-teen-safety-lawsuit-revealed>
- 24 DNA. (n.d.). kouluaitutkimukset vuosilta 2017-2024, viitattu 28.12.2024, <https://corporate.dna.fi/medialle/dnan-tutkimukset/>
- 25 Reuters Institute. (2024). Digital News Report 2024, <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024>
- 26 Reuters Institute. (2024). Digital News Report 2024, <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024>
- 27 Uutismedian liitto. (29.1.2024). Uutisten viikon tutkimus: Tiktok on noussut nuorten tärkeimmäksi uutiskanavaksi, kiinnostus omaan elämään liittyviin uutisiin kasvussa, <https://www.uutismediat.fi/ajankohtaista/uutisten-viikon-tutkimus-tiktok-on-noussut-nuorten-tarkeimmaksi-uutiskanavaksi-kiinnostus-omaan-elamaan-liittyviin-uutisiin-kaavussa/>
- 28 Grönlund, M., Horowitz, M. & Lehtisaari, K. (22.1.2024). Luottamusta rakentamassa, luottamusta kokemassa: Mediapäätäjien ja journalistien yleisöjen näkemyksiä, DECA-hanke, https://cdn.prod.website-files.com/64123799970e0ddca93c317a/65af9e00c156d47a2e021807_Luottamusta%20rakentamassa%2C%20luottamusta%20kokemassa_taitettu_FINAL.pdf
- 29 Meta. (12.3.2021). Learning from videos to understand the world, <https://ai.meta.com/blog/learning-from-videos-to-understand-the-world/>
- 30 Meta. (13.12.2024). Instagram Reels Chaining -tekoälyjärjestelmä, <https://transparency.meta.com/features/explaining-ranking/ig-reels-chaining/>
- 31 Meta. (2024). Facebook Transparency Center, Widely viewed content report: What People See on Facebook, Q3 2024, viitattu 11.12.2024, <https://transparency.meta.com/fi-fi/data/widely-viewed-content-report/>
- 32 Taylor, S. H. & ja Choi, M. (2022). An Initial Conceptualization of Algorithm Responsiveness: Comparing Perceptions of Algorithms Across Social Media Platforms, <https://journals.sagepub.com/doi/10.1177/20563051221144322>
- 33 Stone, S. (19.10.2023). Mastering Social Media Algorithms: The Interest Graph vs the Social Graph, <https://www.theshelf.com/the-blog/social-media-algorithms-interest-graph-vs-social-graph/>
- 34 CMS Law.Tax. (n.d.). GDPR Enforcement Tracker, viitattu 30.12.2024, <https://www.enforcementtracker.com/>
- 35 Euroopan komissio. (n.d.). EU:n digipalvelusäädös, viitattu 30.12.2024, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_fi
- 36 Sitra. (3.4.2024). How to turn off recommendation algorithms on the main social media platforms, <https://www.sitra.fi/en/articles/how-to-turn-off-recommendation-algorithms-on-the-main-social-media-platforms/>
- 37 Pönkä, H. & Mikko, S. (28.10.2024). Mitä hyötyä digipalvelusäädöksestä on diginatiiville? Tarkastelussa Tiktok, Faktabaari

16. AI - threat or opportunity for democracy

MINNA ASLAMA HOROWITZ, UNIVERSITY OF HELSINKI

THREAT OR OPPORTUNITY? The answer is both for generative AI, and for other types and applications of AI. This was the view of Melvin Kranzberg, a professor of technology history, as far back as the 1980s, who even gave his name to Kranzberg's Law: no technology is good or bad, but neither is it neutral. It all depends on how, in what context and for what purpose the technology is used¹.

This is good to remember when faced with either wild threats or glowing reviews of AI. GenAI itself agrees.

When the generative AI application ChatGPT is asked that headline question, it answers as follows:

ChatGPT reminds us that not everyone has access to AI, let alone the ability to use it. AI can also quickly produce propaganda and disinformation and manipulate our opinions in many ways. We also do not know the principles by which different AI applications operate.

On the other hand, ChatGPT can list many good things about itself. AI makes it easier for citizens to become informed and active. AI supports decision-making, for example, by developing scenarios or simplifying complex texts. Generative AI can be used as tools to detect and combat online disinformation. Generative AI can also capture and provide unbiased and balanced perspectives on controversial issues, contributing to public debate.

The open version of ChatGPT produced the above answers in English and Finnish in no time. It's excellent for democratic communication when information can quickly be translated into different languages.

On the other hand, in the summary of threats and opportunities there is not much detail about the effects on the so-called value chain. Firstly, when designing an application, one can consider how it will affect democracy; for example, the fundamental rights of citizens. Secondly, one should consider how data is collected, the quality of the data and the training of the data. Is it done ethically and democratically or not? And finally, the way an application is used can either support or undermine democracy.

ChatGPT's answers, while correct in themselves, illustrate the main challenges and opportunities for democracy presented by generative

AI. Virtually every positive feature ChatGPT lists is true, but the same feature can also be a problem.

For example, if AI is good for us, it is also an excellent tool for the producers of lies, propaganda, and hate speech. Equally, since we often cannot get information on how AI works and what data it uses, we cannot have complete confidence in how it will advise us, or even medical diagnoses or policymakers. So Kranzberg's law can be applied to AI: there are challenges and opportunities.

The answers also show how narrowly ChatGPT understands the big, complex concepts and practices we call democracy. For example, it neatly bypasses privacy and profiling with, say, facial recognition technology. Similarly, it does not mention the economics of AI (i.e. big business) nor the inequalities involved e.g. we can use ChatGPT's free app, but paying for it would give us a better service².

Democracy is about participation in society through work wherever possible, as recalled in the UN Universal Declaration of Human Rights³. However, ChatGPT does not list disappearing jobs or concerns about copyright⁴, as challenges to democracy in its responses described above. Nor does AI mention issues related to sustainable development, particularly energy consumption, despite being an unprecedented energy guzzler⁵. Environmental crises are arguably central to modern societies – and democracy.

With AI already ubiquitous and increasingly public, dozens of statements and regulations exist on ethical and democratic principles. Most of them fear risks – to workers, privacy, national security, and economic growth – but fewer mention the rights and opportunities of ordinary people.⁶

However, if another technology thinker is to be believed, every new technology changes

our relationship with ourselves and each other. More than half a century ago, Marshall McLuhan noted that stories around the campfire built and maintained communities. The world became more individualistic when books became more accessible, since one could read them by oneself⁷. According to McLuhan, the age of electronic communication gave birth to new types of communities where the same stories travel across geographical boundaries. Social media is a prime example of this.

Will generative AI bring us entirely new kinds of communities and societies? The greatest threat to democracy is that we will live with customised and always sympathetic AI-created virtual friends without trusting any of the information we receive – let alone each other – in real life. It is, therefore, time to insist that we all develop AI skills, understand the societal implications of AI and have a say in how, where, why, and who uses generative AI⁸.

¹ Melvin Kranzberg (1986) Technology and History: "Kranzberg's Laws". Technology and Culture 2Vol. 27, No. 3 (Jul, 1986). <https://doi.org/10.2307/3105385>

² E.g. <https://openai.com/chatgpt/pricing/>.

³ UN Universal Declaration of Human Rights <https://www.un.org/en/about-us/universal-declaration-of-human-rights>; <https://um.fi/yk-n-ihmisoikeusjulistus>;

⁴ Generative AI Has an Intellectual Property Problem, <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>; Tekoäly ja EU – myös tekoälyn on kunnioitettava tekijöiden oikeuksia. <https://www.teosto.fi/teostory/tekoaly-ja-eu-myos-tekoalyn-on-kunnioitettava-tekijoiden-oikeuksia/> ;

⁵ AI has an environmental problem, <https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about>; Googlen päästöt karkasivat käsistä tekoälyn vuoksi, <https://yle.fi/a/74-20098239>

⁶ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>; World Economic Forum In The Presidio Recommendations <https://www.weforum.org/publications/the-presidio-recommendations-on-responsible-generative-ai/> and Governing AI for Humanity: Final Report by United Nations https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

⁷ Marshall McLuhan (1964/1994 3rd ed.) Understanding Media. The Extensions of Man. <https://mitpress.mit.edu/9780262631594/understanding-media/>

⁸ Kirsi Hantula (tulossa 2025) Kuinka turvata tiedollisia oikeuksia tekoälyn aikakaudella? SITRA; also e.g., Minna Horowitz, Hannu Nieminen, Katja Lehtisaari & Alessandro D'Arma (2024) Introduction. Epistemic Rights in the Era of Digital Disruption. Palgrave MacMillan. <https://link.springer.com/book/10.1007/978-3-031-45976-4>

17. AI and creativity – Risk or opportunity?

ARI HAASIO, SEINÄJOKI UNIVERSITY OF APPLIED SCIENCES

CREATIVITY IS A COMPLEX concept for which there is no single unambiguous definition. In general, creativity can be thought of as something new that matters!. Creativity is related to both science and art, but it can also be part of business, for example in the form of new, innovative ways of working or the ability to solve problems in a new way. It could be said that creativity is the discovery or invention of a new perspective.

What is the contribution of AI, and can it be creative? Or is it that AI is even stifling our creativity? Current AI applications are all examples of weak AI. Weak or narrow AI can only do what humans tell it to do, while strong AI can think for itself, just like humans. So far, strong AI has not yet been successfully developed².

When considering the role and potential of AI as a facilitator of creative work, it is important to distinguish whether by creativity we mean the instrumental use of AI in the creative process or the end results produced by AI, such as the image it creates.

When discussing the benefits and drawbacks of AI, the focus is always on generative AI, which can produce content such as text and images. It is worth remembering that generative AI is only one form of AI. Different applications of AI are part of our everyday life: a car lane-keeper or a robot vacuum cleaner are examples. Medicine also uses AI, for example in surgical operations³.

AI's own opinion

Let's first ask ChatGPT for its opinion on how it can contribute to the creative process.

AH: How can AI contribute to creativity?

ChatGPT:

Providing inspiration: AI can analyse large amounts of data, such as artwork, music or books, and suggest new ideas or themes that can inspire creative professionals.

Acting as a tool: AI can assist in the creative process, for example in the creation of graphics, music or text. It can automatically generate sketches or melodies that artists can further develop.

Collaboration: AI can be a partner with whom creative people can interact. This collaboration can lead to new perspectives and innovation.

In its response, AI emphasises its own instrumental value – it is not an artist or a researcher, but a tool that allows a creative person to get new ideas, process large amounts of data and sketch things out.

In fact, ChatGPT is at the heart of the matter. If we use AI judiciously in brainstorming and, for example, in sketching and structuring our own projects, it can be a good helper. But we shouldn't let the errand boy be the master. There is a danger in that, even if the temptation is great.

In praise of laziness?

Is AI making us lazier and stupid? Are we no longer working our grey brain cells to solve a problem, acquire knowledge or create new content?

In learning, for example, AI can become a problem: when a student has their tasks done by AI, learning outcomes deteriorate. Problem-based learning is an example of harnessing your own creativity in learning. Already, the breakthrough of ChatGPT has led students to rely on AI. For example, AI is used to outline theses at universities of applied sciences, and an increasing number of high school students are using AI applications to complete translation assignments. In most cases, this is a disservice to personal learning and creative thinking.

The classic Zipf's law, the principle of least effort, is the incentive to use AI: get the information where it's easiest to get it.⁴

Ideas and views?

If AI can be a lazy person's solution for completing a task that requires creative thinking and problem-solving, it is not necessarily always a bad thing. It has various good features, such as the ability to analyse large data sets; which allows humans to focus on cognitive processes rather

than routine activities.

A key part of the creative process when writing, researching or studying a book is data mining.

To facilitate data retrieval, AI applications have been developed for this purpose, such as Iris.AI, which is being tested at the University of Eastern Finland and Haaga-Helia University of Applied Sciences.⁵

The test team was a little sceptical about Iris. It found some useful sources, but information seekers questioned the coverage of the search results. Was there something essential that the informatician had missed?

Artificial intelligence and literature

Artificial intelligence has given artists the tools and opportunities to create new kinds of art. Good examples are the various types of massive video sculptures. Generative AI, on the other hand, has changed the process of making text. It has primarily affected the writing of essays and various types of research papers, but AI-generated fiction and non-fiction is already a thing of the past. In the future, their number is likely to grow significantly.

Artificial intelligence can be used to write entire novels. One of the best-known examples is *Tōkyō-to Dōjō Tō* (Tokyo Tower of Sympathy), by Japanese author Rie Kudo, which won the Akutagawa Prize for Literature in 2023⁶. There are also English-language applications on the market that are specifically designed to produce fiction text. A human defines the genre and some other background variables, and AI writes the script. In my view, this use of AI is detrimental to literature. It easily produces light, formulaic entertainment. These works fail to capture emotion and analytical insight.

Another example is Sudowrite⁷, which makes it easy to write a novel: choose a genre, enter a subject and Sudowrite will write 300 words for you. It suggests a sequel and you can choose from the suggestions or give your own options. In addition,

Writing is a writer's job and should remain so.

Sudowrite allows you to choose a style by giving a narrative style from an author. For example, you can ask the AI to write in the same style as Agatha Christie. In this case, it will formulate the sentences in a “Christie-like” way.

Is writing a book with Sudowrite⁸ or a similar application truly creativity? Does it produce a multi-level, personal and reflective text? Even if it is entertainment, I believe that current AI applications should be used at most to sketch out a topic. Writing is the writer’s job and should remain so.

So far, only a few collections of poems written by AI have been published in Finnish. In addition, AI has been used in the creation of some non-fiction books, mainly as a tool for structuring and brainstorming. In contrast, fiction written by AI has been published in Swedish and English.⁹

Artist’s tool

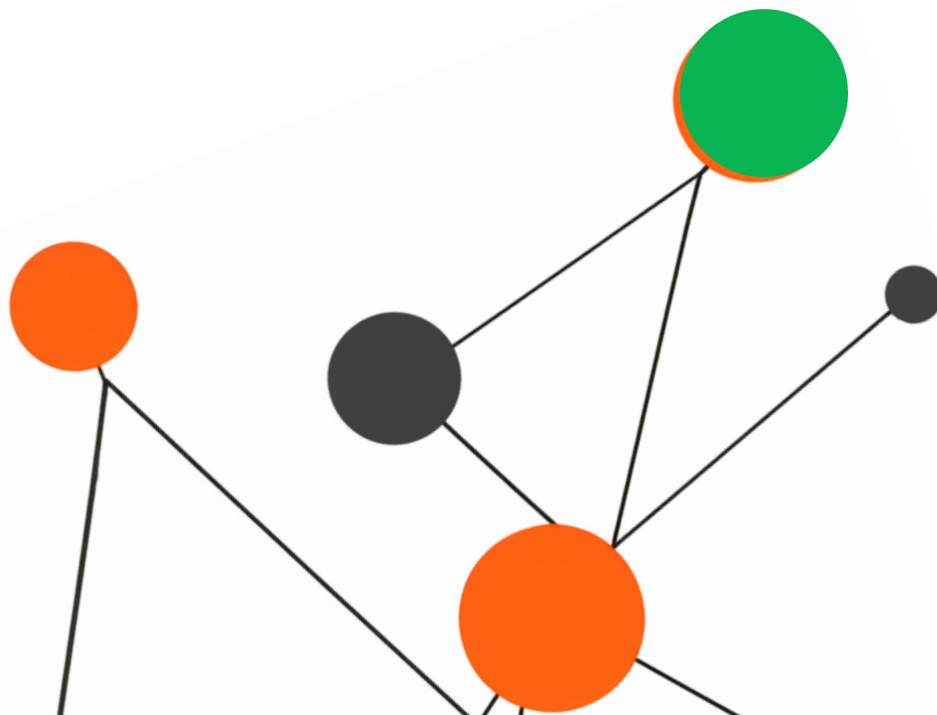
Artificial intelligence has been the subject of heated debate in art circles. Some have even questioned whether art produced by AI is real art¹⁰. Just as Sudowrite turns hobbyists into writers, AI also makes it possible to create images. The problem is often copyright; the work of professional artists is often exploited without permission.

Again, it is a question of how AI is used. For example, Pasi Rauhala has used AI in his

production. It has enabled many solutions that would otherwise have been difficult to implement. An example of this is Rauhala’s Future Truss video sculpture at Aalto University¹¹.

AI certainly enables a layperson to produce images of higher quality than clip art, but ultimately it is humans, not AI, who make the art. Whatever the creative process, AI is a tool that humans should use responsibly. Creativity comes from the human being, and IT adds new dimensions and creates new possibilities. It does not make you an independent artist or researcher, at least for the time being.

-
- 1 Sternberg R. J. (2001). What is the common thread of creativity? Its dialectical relation to intelligence and wisdom. *American Psychologist* 2001; 56. s. 360–362.
 - 2 Haasio, A. (2025). Tekoäly ja kirjastot. *Avain*. (1/2025).
 - 3 Chatterjee, S., Das, S., Ganguly, K., & Mandal, D. (2024). Advancements in robotic surgery: innovations, challenges and prospects. *Journal of Robotic Surgery*, 18 (1), 28.
 - 4 Haasio, A., Harviainen, J. T. & Savolainen, R. (2019). Johdatus tiedonhankintatutkimukseen. *Avain*.
 - 5 Haasio, 2025.
 - 6 Yao, D. (5.2.2024). Author Admits tl. o Using AI to Write Award-winning Novel. *AI Business* 5.2024. <https://aibusiness.com/nlp/author-reveals-using-chatgpt-for-award-winning-novel>
 - 7 Sudowrite. <https://www.sudowrite.com>.
 - 8 Sudowrite. <https://www.sudowrite.com>.
 - 9 Haasio, 2025.
 - 10 Mineo, L. (2023). Writer, animator, architect, musician, and mixed-media artist detail potential value, limit of works produced by AI. *The Harvard Gazette* 15.8.2023. <https://news.harvard.edu/gazette/story/2023/08/is-art-generated-by-artificial-intelligence-real-art/>
 - 11 Salonen, P. (2023). Tekoälyynnii taiteeseen. *Taide* 3/2023.



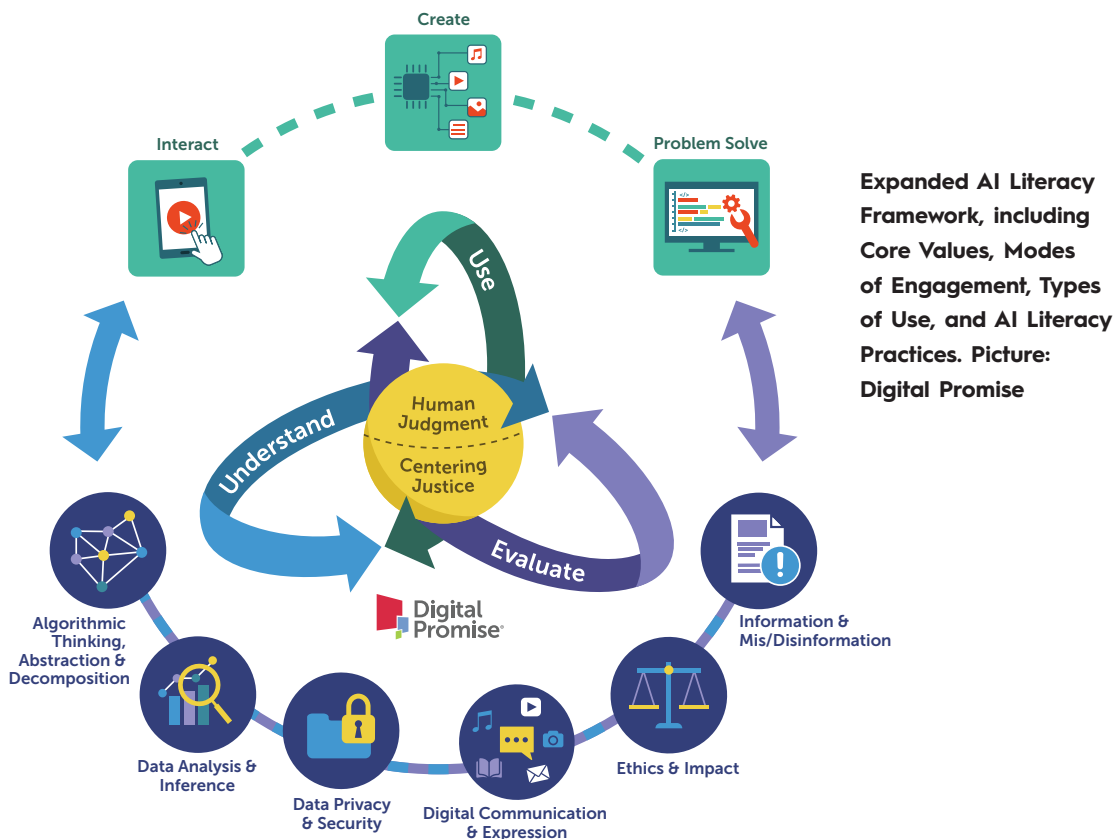
18. Crafting responsible use policies for emerging technologies in US school districts

PATI RUIZ, SENIOR DIRECTOR OF EDTECH AND EMERGING TECHNOLOGIES, DIGITAL PROMISE

HUNDREDS OF APPLICATIONS that integrate GenAI are emerging daily but lack evidence to support claims about their impact and utility for learning, which can affect the appropriate adaptation of policy¹. Users are also rapidly adopting using GenAI tools: The Walton Family Foundation (2023) recently surveyed 1,002 K-12 educators and 1,000 learners in the US and found that 51% of educators are using ChatGPT. Many of them (40%) reported using it at least once a week. Despite its rapid adoption, many people are uncertain about what GenAI is, how it works, and what the implications are for learners, families, and educators.

There are widespread ethical concerns around the adoption of GenAI tools in education² as well as concerns about the environmental impact of these tools³. More work also needs to be done to ensure that AI tools and systems are meeting the accessibility needs of all learners and educators.

In order to understand the integration of GenAI in schools and districts across the United States, and the implications for learners, educators, and families, the GenAI Working Group project convened a diverse



group of education leaders to collaboratively make sense of GenAI in education and inform the development of exemplar R/AUPs that centre responsibility, ethics, and effectiveness.

Research Activities

Digital Promise's League of Innovative Schools (League) network brings together

forward-thinking education leaders from 150 school districts in 38 states. The League is at the forefront of designing, piloting, and scaling solutions that can impact learning outcomes, with a focus on historically and systemically excluded learners. Recent network discussions have highlighted challenges that League leaders are facing regarding AI and emerging technologies for teaching and learning. The project team invited League districts to participate in the GenAI Working Group.

The GenAI Working Group research team and members met once per month in the 2023-

2024 school year to discuss district progress on their *Responsible, Ethical, and Effective Responsible/Acceptable Use Policies (R/AUPs)* and related ongoing conversations they were having with their district communities.

The GenAI Working Group team conducted two focus groups over the course of the project.

Outcomes

The focus groups, interviews, and monthly meetings highlighted the shifting priorities of the participating school districts based on the emergence of AI and emerging technologies in education. Although all GenAI Working Group members were developing strategies to better understand GenAI use in their districts, they understood that their communities' modes of entry differed and that it was important to not rush into implementation.

GenAI Working Group members shared that educators are hesitant to integrate GenAI for several reasons, including the unknown impact of AI on the learning process, overall cost, and in some cases, a lack of AI literacy and technical

expertise. In their conversations with district leaders, learners showed a mixed understanding of how generative AI works, and uncertainty about how to use it productively. Community members shared how their workforces are using generative AI, but had concerns about how rapidly-changing technology will affect both their own daily lives as well as those of learners in their communities.

GenAI Guidance Topics

Conversations among district leaders and their community members reiterated the importance of resources to help understand and integrate AI and other emerging technologies in educational environments. Based on these conversations with districts and grounded in Digital Promise's AI Literacy Framework⁴⁵ and the EdSafe AI Alliance SAFE Benchmarks⁶ we have determined six focus topics for R/AUPs: AI Literacy, Safety, Ethics, Transparency, Implementation Guidelines, and Evaluation and Impact. These topics are described in more detail below.

AI Literacy

This topic provides guidance on how to develop media literacy, digital literacy, and critical thinking skills to understand the safest and most ethical use of AI for learners, educators, and

districts. Community needs in this area were common. We define AI literacy as the knowledge and skills that enable humans to critically understand, evaluate, and use AI systems and tools to safely and ethically participate in an increasingly digital world⁷.

Although AI literacy skills do not need to be explicitly stated in use policies, district leaders can offer guided language on "look-fors" when using GenAI tools. Additionally, rather than developing AI literacy as standalone courses, these skills and practices should be integrated across the curriculum and across grade levels where existing media literacy, digital readiness, and computational thinking skills are also addressed.

Safety

The safety topic prioritises protecting student, teacher, and community data and privacy while managing potential cybersecurity risks, building on federal policies such as the Family Educational Rights and Privacy Act (FERPA)⁸ and the Children's Online Privacy Protection Rule (COPPA)⁹. These policies require parental consent before data is shared and set an age minimum for using AI tools, such as 13 years old for chatbots.

Ethics

GenAI users need to be responsible, fair, and equitable, and acknowledge the biases in both humans and the synthetic outputs of GenAI. Fairness, ethics, and bias have emerged as high priority concerns when using AI. The Working Group shared that educators were primarily concerned about cheating but were also concerned about how to support learners who may be potential victims of bias in AI systems and tools.

Transparency

Leaders should be open about the processes they use when selecting GenAI tools for their schools and about the development and changes in their guidance. As the Office of Educational Technology shares, "In education, decision makers will need more than notice—they will need to understand how AI models work in a range of general educational use cases, so they can better anticipate limitations, problems, and risks"¹⁰.

In addition, guiding language should suggest that learners and educators be transparent when using GenAI tools in their work.

Implementation Guidelines

Guiding language on implementing guidelines includes the ways that educators, learners, and

others can use AI tools to the best of their abilities and to generate results that benefit them and their communities. This topic was front of mind for students and educators: both groups wanted clarity on when it is acceptable and not acceptable

to use AI. Although users have agency to make the ultimate decision whether or not to use AI tools, they still need direction, especially when it comes to emerging technologies.

Evaluation and Impact

The final topic indicated the importance of providing guiding language on how users should evaluate AI systems and tools, as well as the impact of their use, on an ongoing basis. Districts, educators, learners, and community members should continue to ask themselves if individual tools are beneficial or causing harm. Guiding language on evaluation should provide scaffolding on which tools a user may use depending on the need and consider the consequences or benefits of their decision. Helpful guidance on the evaluation of AI can be found in: AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology¹¹.

The GenAI Working Group project's primary deliverable Emerging Technology Guidance Language Deck¹² serves as a living document which will update as districts learn more about how to incorporate language about emerging technologies in their acceptable use policies.

1 Lodge, J., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*, 39(1-8). <https://doi.org/10.14742/ajet.8695>

2 Nguyen, A., Ngo, H.N., Hong, Y., Dang, B., Nguyen, B. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technology* 28, 4221-4241. <https://doi.org/10.1007/s10639-022-11316-w>

3 UNESCO. (2023). ChatGPT and Artificial Intelligence in higher education: Quick start guide. United Nations. https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf

4 Ruiz, P., Armstrong, A. L., Karim, S., Shell, A., Singmaster, H., Giang, M. (2024). Emerging Technology Acceptable Use Policy [Google slides]. Digital Promise. <https://bit.ly/DPAUPLanguage>

5 Mills, K., Ruiz, P., Lee, K., Coenraad, M., Fusco, J., Roschelle, J. & Weisgrau, J. (2024, May). AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology. <https://doi.org/10.51388/20.500.12265/218>

6 EDSAFE AI Alliance. (2024). Safe benchmarks framework. EDSAFE AI. <https://www.edsafeai.org/safe>

7 Mills, K., Ruiz, P., Lee, K., Coenraad, M., Fusco, J., Roschelle, J. & Weisgrau, J. (2024, May). AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology. <https://doi.org/10.51388/20.500.12265/218>

8 FERPA <https://studentprivacy.ed.gov/faq/what-ferpa>

9 COPPA <https://www.ftc.gov/legal-library/browse/rules/childrens-online-privacy-protection-rule-coppa>

10 Office of Educational Technology. (2023). Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations, U.S. Department of Education.

11 Mills, K., Ruiz, P., Lee, K., Coenraad, M., Fusco, J., Roschelle, J. & Weisgrau, J. (2024, May). AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology. <https://doi.org/10.51388/20.500.12265/218>

12 Digital Promise Emerging Technology Guidance Language Deck <https://digitalpromise.dspacedirect.org/items/f84dedc7-762c-460e-b9ba-eb4ba28a9b51>

19. GenAI, copyright and teachers

KARI KIVINEN, FAKTABAARI

THE USE OF AI services raises ethical and intellectual property concerns. In order for teachers and students to work confidently, critically and safely with new and emerging technologies, they should have basic knowledge and a common understanding of how GenAI can be used in education, without forgetting copyright issues.

Generative AI services can be used to create text, images, videos and music. However, there are many copyright uncertainties behind the creative outputs of these services. Rapid technological developments result in legislation lagging behind

Training large language models requires huge amounts of text, code, images and data, much of which may be protected by copyright. If

such data is collected without the permission of the rightholder, there is a risk that someone's copyright may be infringed.

The content of the material used to train the popular language models is, at least for the time being, obscure.

The content of the material used to train the popular language models is currently opaque. The EU's AI Act, which comes into force in 2026, will require future general-purpose AI systems to comply with EU copyright law and provide summaries of the content used to train the data.

The near future will show how this will be implemented. As we do not yet know what information has been used to train AI models, it is difficult to assess potential copyright infringements.

Many lawsuits

AI services can produce output that imitates or closely resembles a copyrighted work. Several artists and publishers have filed lawsuits accusing AI platforms of illegal copying and prohibiting the use of their works for data collection and education. The most high-profile of these lawsuits is the New York Times case against OpenAI. In its defence, OpenAI has issued an interesting response:

"Because copyright today covers virtually every sort of human expression – including blog posts, photographs, forum posts, scraps of software code, and government documents – it would be impossible to train today's leading AI models without using copyrighted materials.."¹

The decisions in these cases will shape and update copyright law and practice for years to come. It is still relatively unclear how we can

prevent AI services from infringing copyright, and what sanctions should be applied and by whom.

The main issue for users of AI services is who owns the outputs of AI. Can traditional copyright apply to them, or are new solutions needed? Currently, copyright does not protect AI-generated works with little or no human input.

Preventing data collection

In particular, large publishers and media houses have begun to block the collection of copyrighted content on their websites. However, blocking data collection can reduce the site's visibility in search engines and prevent legitimate search engines from indexing the site's content.

According to Wired², 88% of leading US news agencies block data collection. The Reuters Institute reported³ that in February 2024, nearly half (48%) of leading news sites in ten countries blocked

OpenAI data collectors. This is something of a problem for GenAI users, as quality and up-to-date content is often excluded from training material. AI services have therefore started to license content from different content providers to ensure quality educational content.

Major publishers and media houses have started to block the collection of copyrighted content on their websites.

This positive trend is likely to increase in the near future.

Copyright and education

Generative AI services may collect information on all input and prompts in their databases in order to train language models. The same can happen, for example, when students' work is scanned by plagiarism checkers.

Particular care should be taken with sensitive data, which should not be fed into AI services under any circumstances. It is therefore crucial to choose AI tools that ensure security practices. For example, they should ensure that data from feeds is not stored or reused for further training purposes. It is also important to remember that most student work is copyrighted material. Education providers should therefore ensure that the AI tools they use comply with EU copyright laws.

EUIPO published an interactive infographic on "**Generative AI in Education – Understanding copyright implications**"

1. Educate your students about copyright.
2. Explain that AI-generated content can still infringe on existing copyright if it reproduces copyrighted material.
3. Teach proper citation skills and how to attribute and credit content generated by GenAI.
4. Promote critical thinking to verify the accuracy of the GenAI outputs.
5. Encourage students to create original works and use AI as a tool for inspiration, not to replace their own creativity.
6. Create interactive lessons where students practice looking for, using, and citing AI-generated content properly.

¹ OpenAI (Taken 12.12.2024) OpenAI faces multiple lawsuits over its use of copyrighted articles, books, and art to train its generative AI tools. <https://www.euronews.com/next/2024/01/09/openai-says-its-impossible-to-train-ai-without-copyrighted-materials>

² Wired (2024) Most top news sites block AI bots. Right-wing media welcomes them. <https://www.wired.com/story/most-news-sites-block-ai-bots-right-wing-media-welcomes-them/>

³ Reuters Institute (2024) How many news websites block AI crawlers. <https://reutersinstitute.politics.ox.ac.uk/how-many-news-websites-block-ai-crawlers>

20. Data protection in AI applications

HARTO PÖNKÄ, INNOWISE

IN THE EUROPEAN UNION, the processing of personal data along with data protection in registers is mainly regulated by the EU General Data Protection Regulation (GDPR)¹. Many other legal acts, such as EU regulations and national laws, also contain provisions related to the processing of personal data and the protection of privacy.

For example, the Finnish Basic Education Act does not directly regulate what personal data must be explicitly processed in education, but its obligations lay the groundwork for municipalities' pupil registers

and, more broadly, what personal data can be processed in basic education in general². A similar situation is typical at different levels of education.

The body responsible for organising education must plan and decide what personal data will be processed and how, within the framework of the laws governing the activities.

The body responsible for organising education must plan and decide what personal data will be processed and how, within the framework of the laws governing the activity. Individual teachers must comply with the policies and guidelines of the education provider. Where the purpose of the processing of personal data is the provision of education, the data may not be used for other purposes such as marketing, product development of the software

provider, or training in general AI models. In e-learning environments, AI applications, and other digital platforms, special care must be taken to respect data protection rules to ensure that the processing of personal data complies with the law.

Data protection begins with planning

In education, the controller of personal data is the body responsible for organising the education; for example, the municipality or the private educational institution. Data protection should be considered right from the planning stage.

One of the first issues to be addressed is the legal basis for processing personal data. The most natural and preferred permitted by the GDPR in education is legal obligation, although this is not without its problems³. In the case of commercial or otherwise contractual education, the processing of personal data is based on the contract in question.

An essential part of planning the processing of personal data involves a prior risk assessment and the safeguards decided on that basis. Safeguards can be both technical and organisational. They are primarily aimed at preventing the use of data for unlawful purposes and ensuring the rights of data subjects under GDPR. These include, on the one hand, ensuring that personal data is not leaked to third parties and is not used for purposes other than those related to the provision of education. And on the other hand, ensuring that individuals can exercise their rights (if they so wish) to receive transparent information about the processing of their personal data and to be able review and correct their data.

On digital platforms, security measures include user IDs, login, network encryption, secure data storage and backup, access levels, timely data deletion, anti-malware, security updates and so

on. As applications are used by people, technical safeguards alone are not enough, but organisational measures such as guidelines, staff training and manual

Individuals can receive transparent information about the processing of their personal data.

controls are also needed.

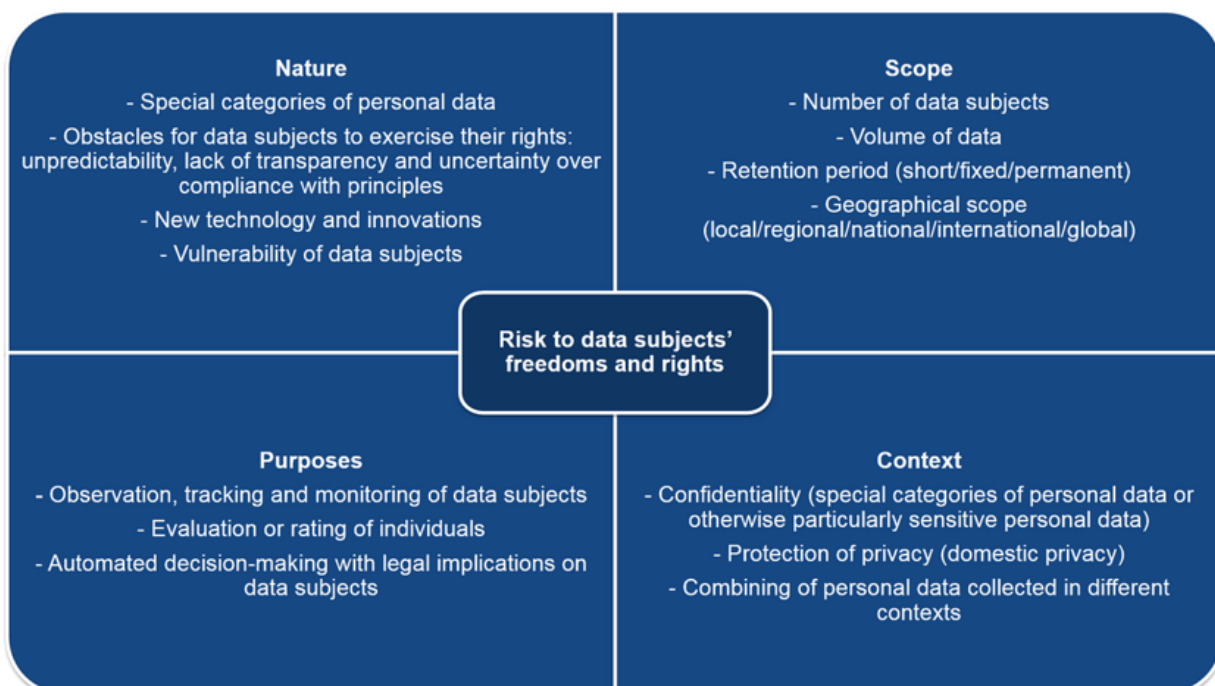
Web services and applications are often complex, both in terms of their actual functionality and their security features. The technical complexity is further increased in AI applications. AI features are also being added to applications where they have not been encountered before.

For example, generative AI functions have been added to conventional image processing software. From a privacy risk assessment perspective, this poses challenges, as the list of functions and features to be assessed can prove to be extensive, even for familiar applications.⁴

AI features are also being added to applications where they have not been encountered before.

Data protection risk assessment of an application

The GDPR requires the controller to carry out a risk assessment of the processing of personal data and implement safeguards to address the risks. The risk assessment will consider both the severity of the potential impact on individuals



and the likelihood of it occurring. If the risk is considered high, the controller shall carry out a data protection impact assessment in accordance with the GDPR. An impact assessment is a procedure provided for by the GDPR that involves a systematic review of the processing of personal data and an identification of threats in order to identify effective safeguards.^{5 6}

Careless use of apps in education can mean a failure to respect data protection and can lead to personal data being compromised. Therefore, AI applications need to be assessed for data protection in the same way as other digital platforms before they are used in education.

Both the education provider as controller and individual teachers can be held liable for potential breaches if they have not followed the employer's instructions on the applications to be used and on data protection considerations. Before implementing AI or other application, teachers should check their organisation's data protection guidelines and policies.

Usually, the provider of the learning platform or application is in the role of processor of personal data, where its liability is limited by contract. A processor in the general data protection regulation means an entity that processes data on behalf of a controller, for example, operating an application because of a service contract.⁷

The risk assessment can be assisted by the following chart, which lists the risk factors related to the nature, scope, context and purposes of the processing.⁸

When it comes to digital platforms and AI applications for teaching, risks are often increased by factors such as:

- **Specific personal data (e.g. health data)**
- **Confidential data (e.g. pedagogical documents and verbal assessments of a person's characteristics)**
- **Vulnerability of the data subject (minors, people with special needs, learners in relation to the education provider)**
- **Location data (to allow systematic monitoring)**

- **Large numbers (e.g. children of compulsory school age in a given municipality)**
- **Long retention period (e.g. nine years of compulsory education)**
- **Confidential data (e.g. data relating to family members and other personal data)**

If the above criteria are met, the situation may be one of high risk for personal data. In addition, an impact assessment must be carried out in several explicitly mentioned cases. These may include, for example, the use of new technologies, the use of location data, the processing of sensitive or very personal data, profiling and automated decision-making which may have a significant impact on the data subject.

In education, it is often the case that the high-risk criteria may be met in some individual cases. For example, specific personal data and confidential information of learners may be stored in the application, even if not intended by the teacher, if learners voluntarily store it in their answers to learning tasks. At the same time, if the application is running an AI-assisted automated review of tasks, this may involve the processing

of personal data with profiling and automated decision-making. The educator must therefore consider the processing of personal data in different situations where the AI features of the application may be combined with the data stored by learners and teachers on the platform. On the other hand, teachers need to be careful not to inadvertently introduce AI features into the application without first assessing them.

In cloud services, which most AI applications are, the server stores many types of personal data, such as user IDs, files and other content, as well as activity and other log data generated using the service. The risk assessment of cloud computing should pay particular attention to

Specific personal data and confidential information of learners may be stored in the application.

the transparency of the processing of personal data, data minimisation and limitation of retention periods, and data confidentiality.¹⁰

To find out how personal data is processed, it is necessary to review the service provider's contracts, security descriptions and other documents. Many software companies have developed their applications primarily for non-EU markets, which may not take into account GDPR requirements. Even if an application advertises GDPR compliance, this does not guarantee that the features meet the level of data protection required for education.

Cloud computing also raises the issue of geographical storage of data and transfers to countries outside the EU and EEA. For example, Microsoft, Google and Adobe web applications

may transfer personal data to countries outside the EU where data protection does not meet the level required by the GDPR. A service contract for a single application can involve up to dozens of sub-

processors in dozens of different countries. For some applications, data storage may be limited to the EU, for example by an additional service or by purchasing a more expensive licence. If this is not possible, other safeguards need to be explored.

Observations on data protection in AI applications

Over the past year, I have been involved in several impact assessments of AI applications and other digital platforms for education. Below, I will discuss my findings. In conclusion, I always recommend careful planning and risk assessment before implementing AI applications and functionalities.

Finding 1: Generative AI applications may lack data protection

Generative AI applications, which have become rapidly popular, are typically free of charge and

their functionalities are targeted at consumers. In the case of educational use, it is worth assessing whether the application is intended for private users or for organisations.

For chatbots based on large language models, the minimum privacy requirement for teaching can be that the input (prompts) written for the application are not used to train that language model. Any personal data provided in the input should not be used for the development of the AI application. If such aspects are not reflected in the documentation of the application, this does not reflect well on the data protection skills of the service provider.

There may also be weaknesses in the security features of applications. For example, if an app allows you to share an AI-generated text or image via a sharing link, such a feature can easily lead to personal data being leaked to an outsider, either accidentally or intentionally, for example for bullying purposes. An educator should be able to monitor the use of the service – including the sharing functionality – to address abuses and remove inappropriate content. It is desirable that overly extensive sharing functions can be switched off.

Shortcomings may even relate to basic data protection issues, such as the possibility to obtain transparent information on the processing of personal data or to verify one's own stored data. At a minimum, viewing or downloading of user data in a computerised format should be available to users at the application administrator level, so that the education provider can fulfil its obligations as a data controller.

Finding 2: Generative AI may have too broad access to stored data

Cloud services for education providers may allow a generative AI such as a chatbot to have access to all files and other data stored in the organisation's cloud service. In such a case, there

Any personal data provided in the input should not be used for the development of the AI application.

It is necessary to review the service provider's contracts, security descriptions and other documents.

is a significant risk that personal data related to education will be used by the generative AI for purposes unrelated to education.

In general, care must be taken over what data is allowed to be accessed by the generative AI or entered as part of a prompt. A good rule of thumb is that the AI should only be given access to individual pieces of data that have been verified not to contain personal data. If AI is to be used to process learners' personal data, the risk assessment described above must be carried out to ensure compliance with data protection regulation.

Finding 3: Generative AI can be offered to student users without a personal data processing agreement

The fact that the same service providers offer both cloud services for educational institutions and generative AI applications open to consumers easily leads to confusion about what is included in the online services offered to educational institutions. A user ID managed by an educational institution may allow access to a generative AI application that is not covered by the contract, in addition to the applications covered by the educational institution contract. Cloud services typically offer a range of licensing options, some of which include and some of which do not include a generative AI application. In this case, it depends on the licence whether the controller of the processing of personal data in the AI application is the education provider or the service provider.

The education provider should check the settings of cloud services when deploying them so that learners' IDs cannot be used to access AI and other web applications whose processing of personal data is not covered by the contract.

Finding 4: The personal data processing agreement is not always comprehensive

Whatever the application, the education provider must check whether a personal data processing agreement can be drawn up with the service

provider or whether it is included in the service contract. If this is not the case, no personal data of learners should be stored in the application without explicitly requested consent for such disclosure and use of personal data. Teachers should check their institution's policy on how to request consent and whether teachers can request consent for the use of applications.

The provider's standard contract may not specify all the personal data to be processed and the purposes of processing. In such cases, it is possible that the data stored in the service may be used, for example, for product development, AI training and other purposes specific to the service provider. If the application contains functionalities that are not covered by the contract, the application provider may be considered as the data controller in this respect.

This would mean that students' data would end up outside the school and for purposes unrelated to education.

The education provider must always check that the contract in force complies with the general data protection regulation and that personal data is processed only for the purposes of the controller. A cautionary example of this is seen in the case of Google Workspaces for Education, whose suitability for educational purposes was questioned by the data protection ombudsman

in 2021. The decision is not yet final¹¹. The Danish data protection authorities, for example, have also pointed out problems with Google cloud services in municipal primary education¹².

The education provider must always check that the contract in force complies with the general data protection regulation.

Guidelines on data protection for the use of AI in education

The use of AI in education must weigh up both the obligations of the data controller responsible for education and the data protection knowledge of teachers, to be able to act correctly in practical situations. The latter requires that teachers' in-

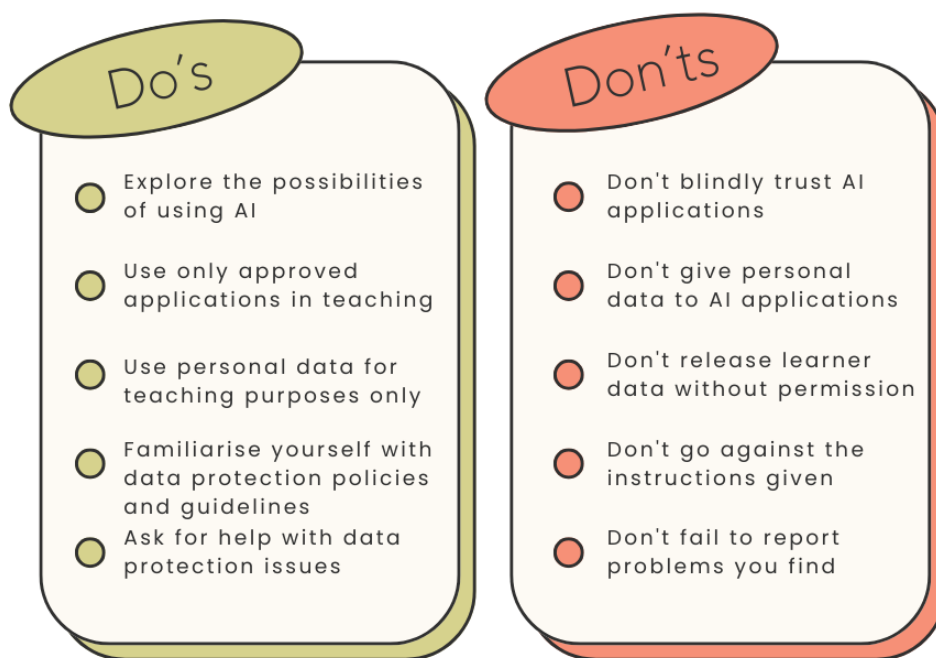
service training covers data protection and the use of AI.

While data protection may seem like a complex issue, addressing it in education starts with simple basics. They stem from the data protection principles at the heart of the general data protection regulation. These include the purpose limitation and minimisation principles, which require that personal data should only be used for the purposes for which it was originally collected and that the data processed should be kept to a minimum. It is worth bearing in

mind the principle of transparency, which can be summarised as the principle that everyone should be able to know what data about them is being processed and how. With these three principles in mind, it is also safe to start using AI applications.

In the following image, I present a set of data protection guidelines for the use of AI in education, which are intended to help teachers consider data protection in practical teaching situations:

Guidelines on data protection for the use of AI in education



¹ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

² Perusopetuslaki 628/1998, <https://www.finlex.fi/fi/laki/ajantasa/1998/19980628>

³ Silvennoinen, E., Tedre, M. & Valttonen, T. (2024). Datafikoituvaa peruskoulu - tasapainoilua lapsen henkilötietojen suojan ja opetuksen tavoitteiden välillä, *Lakimies*, 122(5), 655–678, <https://journal.fi/lakimies/article/view/143755>

⁴ *ibid.*

⁵ Tietosuojatyöryhmä. (2017). Ohjeet tietosuojaa koskevasta vaikutustenarvioinnista ja keinoista selvittää "liittykö käsittelyyn todennäköisesti" asetuksessa (EU) 2016/679 tarkoitettu "korkea riski", <https://tietosuoja.fi/documents/6927448/8316711/Vaikutustenarviointi+fi.pdf>

⁶ Tietosuojavaltuutetun toimisto. (2021). Tietosuojan vaikutustenarvioinnin ohje, <https://tietosuoja.fi/documents/6927448/66036250/TVA+ohje.pdf/ff0b6e1b-5b89-e85e-a2e5-6c4bd4c0ccfc/TVA+ohje.pdf?t=1639729535787>

⁷ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

⁸ Office of the Data Protection Ombudsman. (n.d.). Assess the risks and plan measures to implement data protection, referenced 21.2025, <https://tietosuoja.fi/arvioi-riskit>

⁹ Office of the Data Protection Ombudsman <https://tietosuoja.fi/en/risk-assessment-and-data-protection-planning>

¹⁰ DigiFinland. (26.6.2024). Cirrus-hanke: Tapausesimerkit, <https://digifinland.fi/wp-content/uploads/2024/06/Cirrus-Tapausesimerkit-2024-v1.0.pdf>

¹¹ Tietosuojavaltuutettu. (30.12.2021). Henkilötietojen käsittelyn lainmukaisuus ja siirto kolmansiin maihin koulun opetusohjelman käytössä, <https://www.finlex.fi/fi/viranomaiset/tsv/2021/20211503>

¹² Tietosuojavaltuutetun toimisto. (2.2.2024). Tanskan tietosuojaviranomainen antoi päätöksen Googlen ohjelmistojen käytöstä peruskouluissa, <https://tietosuoja.fi/-/tanskan-tietosuojaviranomainen-antoi-paatoksen-googlen-ohjelmistojen-kaytosta-peruskouluissa>

21. Annex: Instructions for students

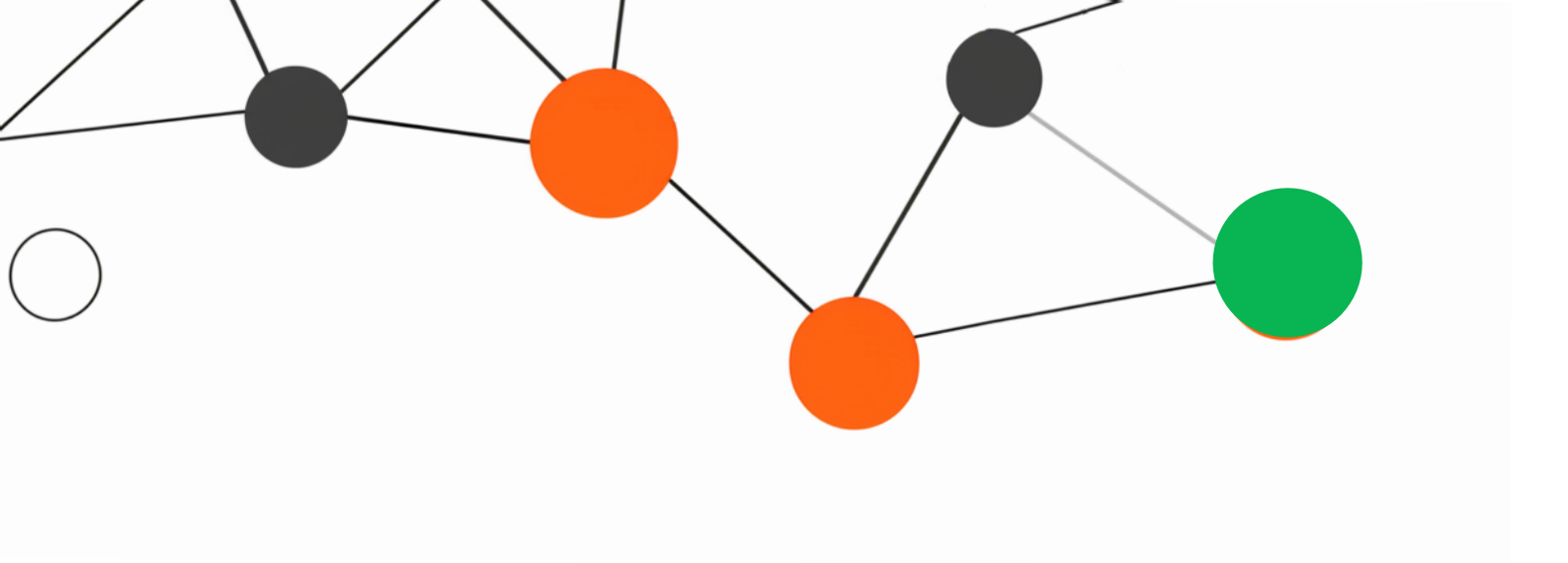
KARI KIVINEN, FAKTABAARI

AI is a tool – use it responsibly!

- **Check the rules:** Find out which AI tools are allowed in your school and in which situations.
- **Be critical:** Remember source criticism: AI can make mistakes. Always check AI's output, the references it provides and compare the information with other sources.
- **Be honest:** Provide sources: openly attribute what has been done with AI in the same way you attribute other sources. Sometimes it is good to share openly what prompts were used so that others can understand the process.
- **Follow instructions:** Follow the instructions given by the teacher.
- **Protect your privacy:** Do not feed confidential or personal information to AI.
- **Be careful:** Follow the safety instructions. There are risks involved in using artificial intelligence. Report any deviations to your teacher.
- **Ask for help:** If in doubt, ask your teacher or friends for advice.
- **Help others:** Share your knowledge with others. AI can be new to many people, and by sharing your knowledge you can help others to learn.
- **Limit:** Use AI applications in moderation so that they don't take up all your time.
- **Act responsibly:** Make sure you don't infringe copyright or plagiarise other people's work. Using AI responsibly also means respecting the work of others and following ethical principles.

Remember:

- AI is not a substitute for your own thinking.
- AI can be a useful tool, but it must be used responsibly.
- AI is not human and cannot think for itself.



READ ALSO: FAKTABAARI'S DIL GUIDE!

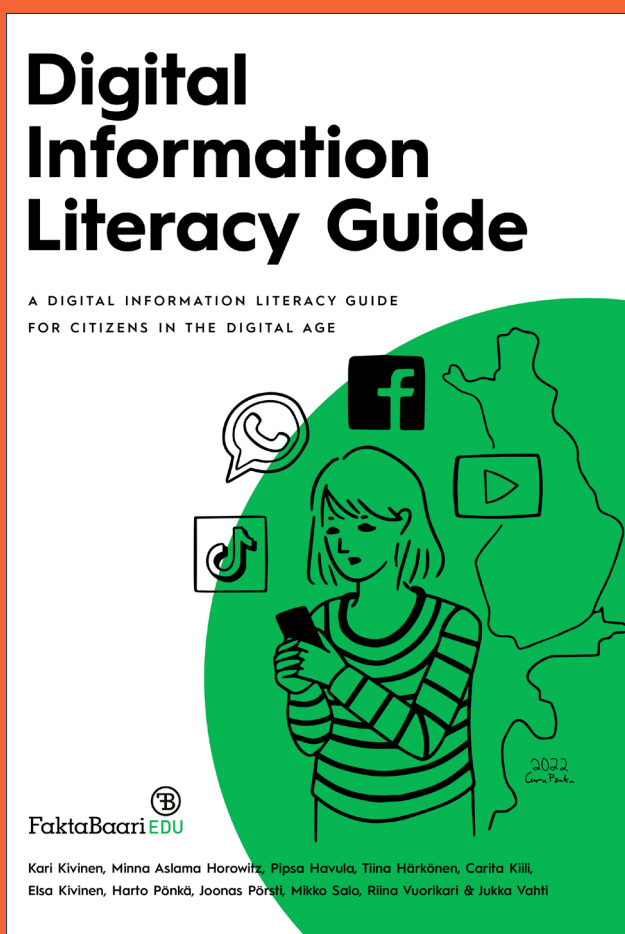
DIGITAL INFORMATION LITERACY GUIDE. A DIGITAL INFORMATION LITERACY GUIDE FOR CITIZENS IN THE DIGITAL AGE

Digital information literacy is a modern civic skill that underpins participation in democratic decision-making. Finland is renowned for its high literacy rate, and the teaching of multiple literacies has been

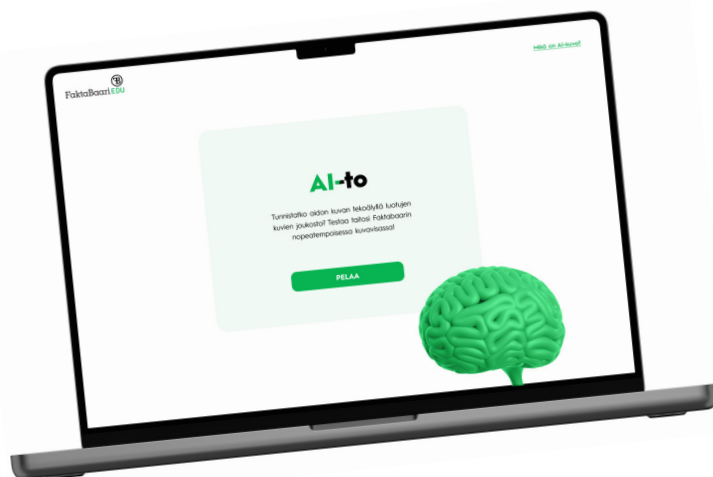
integrated into current curricula from early childhood education onwards.

However, on digital platforms we all are confronted with a bewildering flood of information that we may not be able to filter out with the skills we have acquired in the school community and at home: claims about products by influencers, search results tailored by commercial algorithms, cleverly scripted propaganda and authorisations to track online behaviour or physical movement in urban space hidden behind countless 'yes' buttons. It is therefore important to strengthen the digital information literacy of all the web users, especially young people, in order to identify how we are being influenced online.

This guide is a collection of information and examples on how to enhance digital information literacy.



STOP THINK CHECK



peli.faktabaari.fi