



MINISTÈRE
DE L'ÉDUCATION
NATIONALE
ET DE LA JEUNESSE

*Liberté
Égalité
Fraternité*

Artificial intelligence and education

Research contributions
and challenges
for public policies

January 2024



GTnum

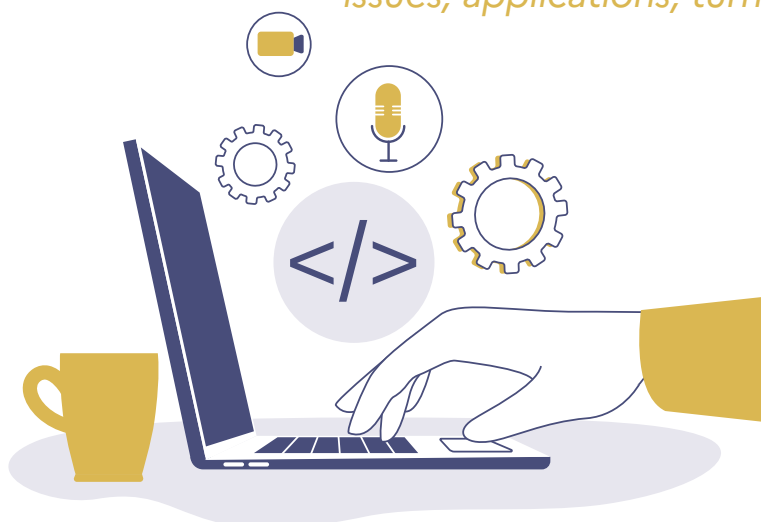
État des lieux sur les apports de la recherche et les lignes directrices des institutions internationales sur l'intelligence artificielle et l'éducation :

définitions et approches, politiques publiques, enjeux éthiques, applications, tournant des IA génératives.



Current situation on the Research Contributions Guidelines of international institutions on artificial intelligence and education:

definitions and proposals, public policies, ethical issues, applications, turning point of Generative AI.



Contents

- Abstract..... 4
- Key points 5
- 1. Artificial intelligence (AI): manifold definitions and approaches 6**
- 2. What are the challenges for public policies? ... 13**
- 3. What are the ethical considerations?..... 23**
- 4. What are the areas of application?..... 25**
- 5. What solutions exist for being trained on and teaching AI?..... 30**
- 6. The tipping point of generative AI systems and large language models 38**
- References..... 50

Abstract

This document, produced by the Digital Directorate for Education (Ministry of National Education and Youth), offers an inventory (January 2024) of the contributions of research and the guidelines of international institutions on the artificial intelligence (AI) and education: diversity of definitions and approaches, issues for public policies, ethical issues, fields of application, avenues for training and teaching, perspectives with the turning point of generative AI and great language models.

Résumé

Ce document, réalisé par la Direction du numérique pour l'éducation / ministère de l'Éducation nationale et de la Jeunesse, propose un état des lieux (janvier 2024) sur les apports de la recherche et les lignes directrices des institutions internationales sur l'intelligence artificielle (IA) et l'éducation : diversité des définitions et des approches, enjeux pour les politiques publiques, enjeux éthiques, domaines d'application, pistes de travail pour former et enseigner, perspectives avec le tournant actuel des systèmes d'IA générative et des grands modèles de langage.

Resumo

Este documento, producido por el Departamento de Educación Digital (Ministerio de Educación Nacional y Juventud), ofrece un inventario (Enero de 2024) de las contribuciones de la investigación y las directrices de las instituciones internacionales sobre la inteligencia artificial (IA) y la educación: diversidad de definiciones y enfoques, cuestiones para las políticas públicas, cuestiones éticas, campos de aplicación, vías de formación y enseñanza, perspectivas con el giro actual de las IA generativas y los grandes modelos de lenguaje.

Key points

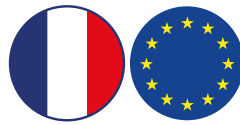
Artificial intelligence: manifold definitions and approaches

"Surprisingly, the term recently used to describe these astounding computing abilities is artificial intelligence (AI). The return of this term – coined by John McCarthy in 1956 – to the front stage is an interesting enigma in the history of science and technology."
(Cardon et al., 2018)



What are the challenges for public policies?

"The AI strategy in education must provide solutions for education and training professionals and help with learning and decision-making without ever replacing people."
(French Government – Directorate General for Enterprise, 2023)



What are the ethical considerations?

The European Commission has put forward four key considerations: human action and human control; fairness; human-centred; justification of choices, transparency and explainability.
Source: (Directorate-General for Education, Youth, Sport and Culture & European Commission, 2022)



Areas of application in education

"Artificial intelligence plays an important and growing role in education. One important case is of personalised teaching systems which are already well established, with growing evidence of their effectiveness for improving learning."
(Khosravi et al., 2022)



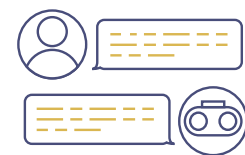
What solutions exist for being trained on and teaching AI?

A proposed five pillar pathway
1. Uncertainty and randomness
2. Coding and calculation
3. Data awareness
4. Critical thinking
5. Post-AI humanism
Source: (de la Higuera & Bocquet, 2020)



The tipping point of generative AI systems and large language models

"During the last two years there has been a plethora of large generative models such as ChatGPT or Stable Diffusion that have been published. Concretely, these models are able to perform tasks such as being a general question and answering system or automatically creating artistic images that are revolutionizing several sectors."
(Gozalo-Brizuela & Garrido-Merchan, 2023)



1 Artificial intelligence (AI): manifold definitions and approaches

1.1 Origins and founding acts

Going back to its origins, as a new scientific discipline, AI was “based on the conjecture that all **cognitive functions** – especially learning, reasoning, computation, perception, memorization, and even scientific discovery or artistic creativity – can be described with such precision that it would be **possible to programme a computer to reproduce them.**”
Source: (Ganascia, 2022)

“In addition to an important paper on morphogenesis in 1952, [Turing’s] two most famous papers in the world of computing are ‘On Computable Numbers, With an Application to the Entscheidungsproblem’ (1937), where he laid the foundations of **Turing machines** and computability theory, and ‘Computing Machinery and Intelligence’ (1950), where he posited the **question of whether machines could exhibit intelligent behaviour** and proposed the Turing test.” (Le Blanc, 2014)

“One classical definition specifies that AI is research that develops technologies that can do things that would require intelligence if done by humans (Minsky, 1969). This approach originates from **Turing**, who proposed that if a **simulation of an intelligent human** cannot be distinguished from a real person, questions about intelligence become irrelevant (Turing, 1950). Many cognitive scientists and some AI researchers and philosophers have adopted a stronger view, arguing that research on AI can reveal how the human mind works (Gardner, 1985).” (Holmes et al., 2022)

“If its history needs a beginning, let’s start with the **Dartmouth Conference**, where the term ‘artificial intelligence’ was first coined. It was held during the summer of **1956** at Dartmouth College near Hanover, New Hampshire, by AI’s two pioneers: **Marvin Minsky** and John McCarthy.” (Le Cun et al., 2019)

“Surprisingly, the term recently used to describe these astounding computing abilities is *artificial intelligence* (AI). The return of this term – **coined by John McCarthy in 1956** – to the front stage is an interesting enigma in the history of science and technology. Specifically, the majority of close observers state that it is only in the field of machine learning methods, and in particular in deep learning, that observable progress in calculated prediction is currently taking place. Yet these techniques have not always been considered to fall within AI. (...)

The origins of **neural networks** are found in the pioneering history of computer science and early cybernetics. Even though the term was coined later, cybernetics can effectively be considered 'connectionist' and still refers to the goal of mathematically modelling a neural network, set by neurophysiologist **Warren McCulloch** and logician **Walter Pitts in 1943**. To this day, that seminal article continues to be quoted as the starting point of the connectionist journey, even in current citations in deep learning articles." (Cardon et al., 2018)

Milestones in the history of AI

| | |
|-----------|---|
| 1943 | McCulloch and Pitt publish their article on artificial neurons . |
| 1950 | Turing publishes his article "Computing Machinery and Intelligence" based on the Turing test. |
| 1956 | The term "artificial intelligence" is used for the first time at a conference held at Dartmouth College . |
| 1958 | John McCarthy invents the programming language LISP. |
| 1964-1966 | ELIZA , one of the first ever chatbots , is developed by Joseph Weizenbaum. It gives rise to the "Eliza effect", the tendency to unintentionally project human traits into sequences of computer-generated symbols. |
| 1972-1981 | The MYCIN expert system is developed, helping physicians to diagnose infectious diseases. |
| 1974 | Paul Werbos develops the first backpropagation algorithm, a key component in deep learning . |
| 1997 | Deep Blue , the chess computer developed by IBM, beats chess world champion Garry Kasparov. |
| 2011 | IBM's Watson wins the TV game "Jeopardy!" against two former human champions. |
| 2012 | The AlexNet convolutional neural network wins the ImageNet competition, marking the beginning of the deep learning era. |
| 2016 | AlphaGo , an AI program developed by Google DeepMind, beats world Go champion Lee Sedol. Google develops its voice assistant. |
| 2017 | Article on Transformers published by (Vaswani et al.) [Google Research]. |
| 2018 | Google develops its BERT language model , improving performance in automatic natural language processing. |
| 2019 | Pluribus developed by Facebook AI makes a breakthrough in the field of poker. |
| 2022 | OpenAI launches ChatGPT to the public. |
| 2023 | Rapid development of generative AI systems and large multimodal models (text, image, sound, video). |

1.2 Human decisions, modelling and automation of cognitive tasks

“The term ‘artificial intelligence’, coined in the 1950s, refers to the **field of research that studies the mechanisms of intelligence** by modelling them using algorithms and experimenting them with machines. These mechanisms include, for example, the ability to automatically find solutions to problems, which can then implement planning, prediction, control, memorisation or learning functions. By extension, the term ‘artificial intelligence’ is often used to refer to **algorithms** [set of rules followed by calculations for problem-solving] **that simulate or share characteristics with certain human intelligence capabilities.**” (Inria & Class’Code, 2020)



AI is defined as “**machine-based** systems that can, given a set of human-defined objectives, **make predictions, recommendations, or decisions** that influence real or virtual environments. AI systems interact with us and act on our environment, either directly or indirectly. Often, they appear to operate autonomously, and can adapt their behaviour by learning about the context.” Source: (UNICEF, 2021) and cited by (Holmes & Tuomi, 2022)

It is also a “**theoretical and practical interdisciplinary field** that combines digital, mathematical, statistical and algorithmic technologies.” (*Grand lexique français de l’Intelligence artificielle*, 2021)

“Thus, we can define artificial intelligence (AI) as **computing systems that are able to engage in human-like processes** such as learning, adapting, synthesizing, self-correction and use of data for complex processing tasks.” (Popenici & Kerr, 2017) “[This definition] allows us to **overcome the opposition between human intelligence and AI and consider complex interactions between the two.**” (Romero *et al.*, 2023)

1.3 Types of AI

There are several types of AI:

Defined by their technology

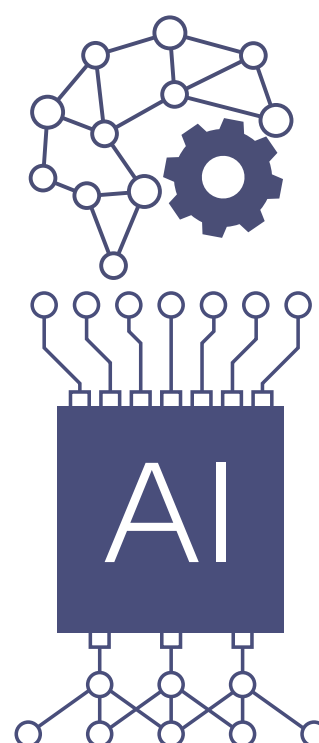
The other is knowledge-based or symbolic AI. **Data-driven AI** might have great potential in education, depending on the intentions of the system, but **knowledge-based AI** still underpins most existing AIED systems. A third conceptual model, **hybrid AI** that combines data-driven and knowledge-based approaches with human cognition." Source: (Holmes et al., 2022)

Defined by their scope

"**Weak AI differs from strong AI** in terms of the scope of application of these computer programs. Weak AI is a type of artificial intelligence that focuses on a specific task, while strong AI or AGI (artificial general intelligence) can be applied to any problem. All existing AI systems are currently considered weak artificial intelligence systems." (*Grand lexique français de l'Intelligence artificielle*, 2021)

"**Complete or strong artificial intelligence** [General AI]: Artificial intelligence that has the ability to mimic human functions (learning, understanding, reasoning, decision-making, having a conscience, feeling emotions, etc.). **To date, strong artificial intelligence does not exist: it is a concept.**

Weak artificial intelligence: This is artificial intelligence as we know it today: it's a 'learning' algorithm that adapts its parameters to learn data and is not equipped with mental or cognitive functions, but can perform a specific task efficiently, sometimes more efficiently than humans." (Inria & Class'Code, 2020)



Classic AI and machine learning

Classical AI: "Much early or 'classical AI', variously known as '**symbolic AI**', 'rule-based AI', or 'good-old-fashioned AI' ('GOFAI'), involves writing sequences of IF... THEN... and other rules of conditional logic, steps that the computer will take to complete a task. Over decades, rule-based AI '**expert systems**' were developed for a diverse range of applications, such as medical diagnostics, credit ratings, and manufacturing. Expert systems are based on an approach known as 'knowledge engineering', which involves eliciting and modelling the knowledge of experts in a specific domain, a resource-intensive task that is not without complications. (...)

Machine learning: Many recent AI advances – including natural language processing, facial recognition, and self-driving cars – have been made possible by advances in machine-learning-based computational approaches. Rather than using rules, **machine learning (ML) analyses large amounts of data to identify patterns and build a model** which is then used to predict future values. It is in this sense that the algorithms, rather than being pre-programmed, are said to be 'learning'.

There are **three main ML approaches**:

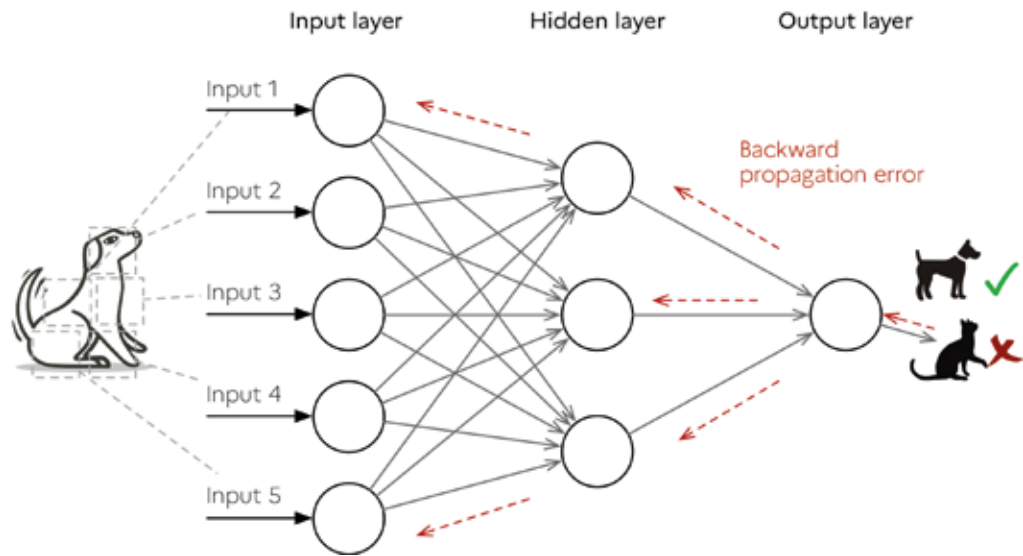
- **supervised,**
 - **unsupervised,**
 - **and reinforcement.**
-
- **Supervised learning** involves data that has already been labelled – such as many thousands of photographs of people that have been labelled by humans. The supervised learning links the data to the labels, to build a model that can be applied to similar data – for example, to automatically identify people in new photographs. In unsupervised learning, the AI is provided with even larger amounts of data, but this time the data has not been categorized or labelled.
 - **The unsupervised learning** aims to uncover hidden patterns in the data, clusters that can be used to classify new data. For example, it may automatically identify letters and numbers in handwriting by looking for patterns in thousands of examples. In both supervised and unsupervised learning, the model derived from the data is fixed, and if the data changes, the analysis has to be undertaken again.
 - However, the third ML approach, **reinforcement learning**, involves continuously improving the model based on feedback – in other words, this is machine learning in the sense that the learning is ongoing. The AI is provided with some initial data from which it derives a model, which is assessed as correct or incorrect and rewarded or punished accordingly. (...)

An artificial neural network (ANN) is an AI approach that is inspired by the structure of biological neural networks (i.e. animal brains). ANNs each comprise three types of interconnected layers of artificial neurons: an input layer, one or more hidden intermediary computational layers, and an output layer that delivers the result (...)

Deep learning refers to ANNs that comprise multiple intermediary layers. It is this approach that has led to many of the recent remarkable applications of AI (for example, in natural language processing, speech recognition, computer vision, image creation, drug discovery, and genomics).“

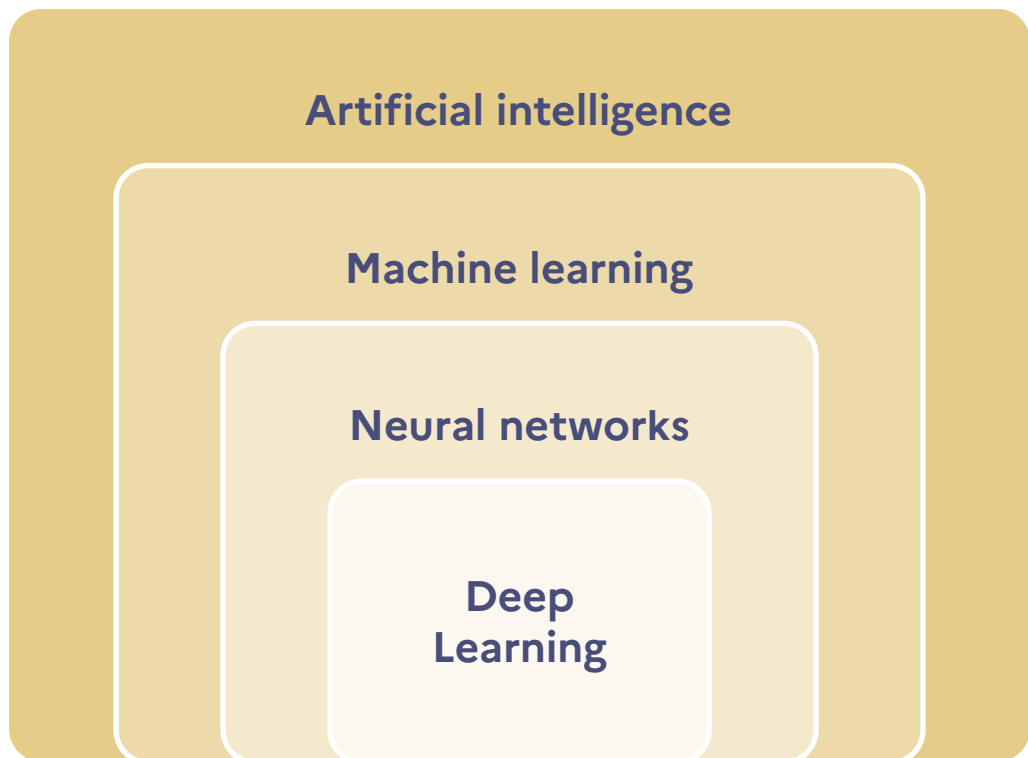
(UNESCO, 2021)

Working of a neural network



(Cardon et al., 2018)

AI, machine learning, neural networks and deep learning



(UNESCO, 2021)

What is Edge AI?

"AI at the edge is the use of AI in real-world devices. Edge AI refers to the practice of doing AI computations near the users at the edge of the network instead of centralised locations like a cloud service provider's data centre or a company's own private data warehouse."

(Singh & Gill, 2023)

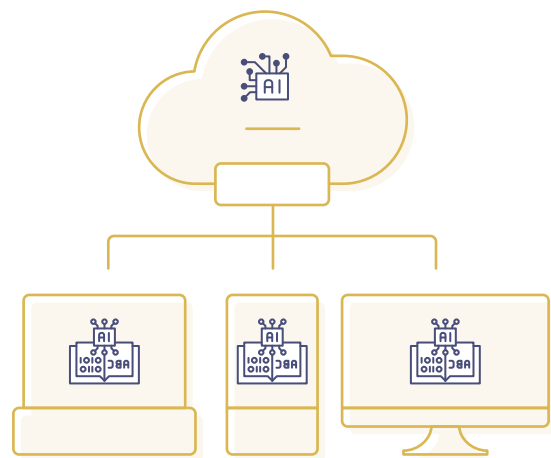
Edge AI

Cloud or centralised AI



Data and computing power centralised in the cloud (remote servers)

Edge or decentralised AI



Data and computing power distributed between remote servers and endpoints

2. What are the challenges for public policies?

2.1 The Beijing Consensus

The Beijing Consensus “recommend[s] that governments and other stakeholders in UNESCO’s Member States:

- **Plan AI in education policies** to leverage the opportunities and tackle the challenges of AI technologies, adopt cross-sectoral and multi-party approaches that engage the entire government and promote local strategic priorities to achieve SDG 4 [Sustainable Development Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all].
- **Support the development of new models** ushered in by AI technologies for the provision of education and training services where the benefits clearly outweigh the risks, and use AI tools to deliver lifelong learning systems that enable personalised learning anytime, anywhere, for all.
- **Envisage the use of pertinent data**, where appropriate, to encourage evidence-based policy planning.
- Ensure AI technologies are used to **empower and not replace teachers** and develop programs designed to build skills to enable teachers to work alongside AI systems.
- Prepare the next generation of workers by **equipping them with the most pertinent values and skills needed for living and working** in the age of AI.
- Promote a **fair and inclusive use** of AI, regardless of disability, social or economic status, ethnic or cultural origin or geographical location, with emphasis on gender equality, while ensuring ethical, transparent and verifiable uses of educational data.”

(UNESCO, 2019)



2.2 UNESCO and OECD recommendations and forecasts

Recommendations for policy makers

1. A system-wide vision and strategic priorities

- Define a system-wide vision of ai and education policies
 - Ensuring the inclusive and equitable use of AI in education;
 - Leveraging AI to enhance education and learning;
 - Promoting the development of skills for life in the age of AI, including teaching how AI works and its implications for humanity;
 - Safeguarding the transparent and auditable use of education data.
- Assessing system-wide readiness and choose strategic priorities

2. Overarching principle for AI and education policies

- Adopt a humanistic approach as an overarching principle for ai and education policies

3. Interdisciplinary planning and inter-sectoral governance

- Mobilize interdisciplinary and multi-stakeholder expertise to inform policy planning and build the capacities of policy-makers
- Set up inter-sectoral governance and coordination mechanisms

4. Policies and regulations for equitable, inclusive, and ethical use of AI

- Set out cross-cutting strategic objectives, and plan regulations and programmes, to ensure the equitable and inclusive use of ai in education

5. Master plans for using AI in education management, teaching, learning, and assessment

- Leverage ai to boost and upgrade education management and delivery
- Cultivate learner-centred use of ai to enhance learning and assessment
- Ensure that ai is used to empower teachers
- Plan the use of ai to support lifelong learning across ages, locations and backgrounds
- Develop values and skills for life and work in the ai era

6. Pilot testing, monitoring and evaluation, and building an evidence base

- Build a trusted evidence base to support the use of ai in education
- Strengthen research and evaluation in the field of ai and education

7. Fostering local AI innovations for education

- Promote the local development of ai technologies for education

Source: (UNESCO, 2021) (Extracts)



The new challenges of generative AI

“The implications of GenAI for assessment go far beyond the immediate concerns about learners cheating on written assignments. We must contend with the fact that **GenAI can produce relatively well-organized papers and essays and impressive works of art**, and can pass some knowledge-based exams in certain subject areas. **We therefore need to rethink what exactly should be learned and to what ends**, and how learning is to be assessed and validated. (...)

The most fundamental perspective of the long-term implications of GenAI for education and research is still about the **complementary relationship between human agency and machines**. One of the key questions is whether humans can possibly cede basic levels of thinking and skill-acquisition processes to AI and rather concentrate on higher-order thinking skills based on the outputs provided by AI.”

(Miao & Holmes, 2023)

“Generative AI models that generate text, image, video, and audio (e.g., music, speech) content are advancing at breakneck speed. This poses **endless possibilities, demonstrated across a growing array of domains**. However, the technology also poses numerous challenges and risks to individuals, companies, economies, societies, and policymaking around the globe, ranging from near-term labour-market disruption and disinformation to potential long-term challenges in controlling machine actions. The future trajectories of generative AI are difficult to predict, but governments must explore them to have a hand in shaping them.

Technological development of generative AI is in its nascent stage, with first-movers such as established tech players like Microsoft, Google and Meta, and private research labs like OpenAI, Midjourney, and Stability.AI. These firms are pursuing multiple strategies to capitalise on generative AI and, to some extent, mitigate its downsides.

Public discussion about generative AI is less than a year old. With technology companies bringing generative AI applications to market, policy makers around the globe are grappling with its implications. Applied and academic researchers are engaged in a fierce debate about how to handle generative AI, from mitigation measures in model design and development, through market launch and beyond.

The path ahead is unclear and replete with differing perspectives. One extreme argues for a moratorium on experiments with generative AI more advanced than GPT-4 (Future of Life Institute, 2023) while the other believes that the supposed existential risks of AI are overhyped (LeCun and Ng, 2023). Others— perhaps most—fall somewhere in between. Regardless of ideological stance on these issues, **there is an urgent need for further research to prepare for different possible generative-AI future scenarios**. Given the great uncertainty and potentially large impact the technology could have at both micro and macro levels, policy makers must remain informed and prepared to take appropriate action through forward-looking AI policies.

The OECD intends for this paper to serve as a steppingstone to help governments make progress in this area. The OECD.AI Policy Observatory and its new OECD Expert Group on AI Futures will serve alongside other relevant bodies as a forum for dialogue on these topics, generating insights and actionable recommendations for governments. Complementary work is also underway through other OECD initiatives, such as work conducted under the Employment, Labour and Social Affairs Committee, the OECD DIS/MIS Resource Hub and the horizontal OECD Going Digital initiative.”

(Lorenz *et al.*, 2023)

2.3 A national strategy

“Artificial intelligence (AI) is a priority for France in the areas of research, economics, modernisation of public action, regulations and ethics. The Government launched a national AI strategy in 2018.

Following the report ‘Giving meaning to artificial intelligence: towards a national and European strategy’ (2018) and as part of the France 2030 plan, the Government launched a **national strategy for artificial intelligence (SNIA)** in 2018. This strategy aims to position France as one of the European and global leaders in artificial intelligence (AI).

The national strategy for artificial intelligence aims to safeguard and strengthen France’s economic, technological and political sovereignty and put AI at the service of the economy and society. It is linked to the mastering sovereign and secure digital technologies component of the France 2030 plan.

The national AI strategy is divided into two phases:

PHASE 1 (2018-2022)

Equipping France with competitive research capabilities. 1.85 billion euro of funding was earmarked for phase 1. In particular, it funded the creation and development of a network of interdisciplinary artificial intelligence institutes, the establishment of chairs of excellence and doctoral programmes, as well as the deployment of the Jean Zay supercomputer.

PHASE 2 (2021-2025)

Rolling out AI technologies across the economy and supporting development and innovation in priority areas such as embedded AI, trustworthy AI, frugal AI and generative AI. This second phase is earmarked with 1.5 billion euro of funding under France 2030. This new phase of the SNIA is built around three strategic pillars: supporting the deep tech offering, training and attracting talent, and bringing together the supply and demand for AI solutions.”

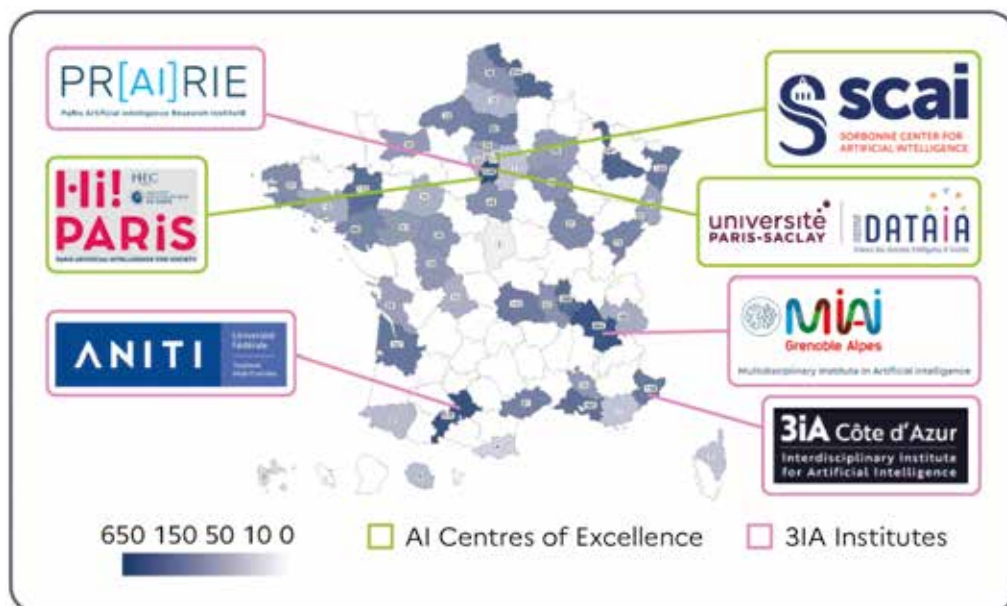
(French Government – Directorate General of Enterprise, 2023)

Timeline and location

Development phases of national strategy in artificial intelligence



Establishment of AI-specialised institutes (3IA PR[AI]RIE, MIAI, 3IA Côte d'Azur and ANITI) and centres of excellence (SCAI, DATAIA and Hi! PARIS) in regions historically active in this field



Source: Cour des comptes

Note to reader: Historical activity in AI is measured through the regional distribution of theses on AI defended between 1989 and 2019 in French higher education institutions based on open-source data from theses.fr (ABES).

Theses are presented by year of defence. AI thesis detection is carried out using the semantic method developed by Cour des Comptes. The logos of the 3IAs and AI Centres of Excellence are taken from their official websites.

(Cour des comptes, 2023)

2.4 A European AI regulation (AI Act)

“MEPs reached a political deal with the Council on a bill to ensure AI in Europe is safe, respects fundamental rights and democracy, while businesses can thrive and expand.

On Friday [08/12/23], Parliament and Council negotiators reached a provisional agreement on the Artificial Intelligence Act. This regulation aims to ensure that fundamental rights, democracy, the rule of law and environmental sustainability are **protected from high risk AI**, while boosting innovation and making Europe a leader in the field. The rules establish obligations for AI based on its potential risks and level of impact. (...)

Banned applications: biometric categorisation systems that use sensitive characteristics (e.g. political, religious, philosophical beliefs, sexual orientation, race); untargeted scraping of facial images from the internet or CCTV footage to create facial recognition databases; emotion recognition in the workplace and educational institutions; social scoring based on social behaviour or personal characteristics; AI systems that manipulate human behaviour to circumvent their free will; AI used to exploit the vulnerabilities of people (due to their age, disability, social or economic situation).

Measures to support innovation and SMEs: MEPs wanted to ensure that businesses, especially SMEs, can develop AI solutions without undue pressure from industry giants controlling the value chain. To this end, the agreement promotes so-called regulatory sandboxes and real-world-testing, established by national authorities to develop and train innovative AI before placement on the market.”

Classification of AI systems

“The compromise agreement provides for a **horizontal layer of protection**, including a high-risk classification, to ensure that AI systems that are not likely to cause serious fundamental rights violations or other significant risks are not captured. AI systems presenting only **limited risk** would be subject to very light transparency obligations, for example disclosing that the content was AI-generated so users can make informed decisions on further use.

A wide range of **high-risk** AI systems would be authorised, but subject to a set of requirements and obligations to gain access to the EU market. These requirements have been clarified and adjusted by the co-legislators in such a way that they are more **technically feasible** and **less burdensome** for stakeholders to comply with, for example as regards the quality of data, or in relation to the technical documentation that should be drawn up by SMEs to demonstrate that their high-risk AI systems comply with the requirements.

Since AI systems are developed and distributed through complex value chains, the compromise agreement includes changes clarifying the **allocation of responsibilities** and **roles of the various actors** in those chains, in particular providers and users of AI systems. It also clarifies the relationship between responsibilities under the AI Act and responsibilities that already exist under other legislation, such as the relevant EU data protection or sectorial legislation.

For some uses of AI, risk is deemed **unacceptable** and, therefore, these systems will be banned from the EU. The provisional agreement bans, for example, cognitive behavioural manipulation, the untargeted scraping of facial images from the internet or CCTV footage, emotion recognition in the workplace and educational institutions, social scoring, biometric categorisation to infer sensitive data, such as sexual orientation or religious beliefs, and some cases of predictive policing for individuals.”

Source: (European Parliament, 2023) and (Council of Europe, 2023)

2.5 Actions and output of digital theme groups (#GTnum)

Actions by GTnum IA and open education



(DNE-TN2, 2023a)

GTnum portfolios on AI



Portfolio details and content: (DNE-TN2, 2023c)



Portfolio details and content: (DNE-TN2, 2023b)

2.6 AI performance in reading and mathematics

“The study is part of a comprehensive project for assessing AI capabilities and their implications for work and education, led by the OECD Centre for Educational Research and Innovation (CERI). The **AI and the Future of Skills (AIFS) project** aims to develop measures of AI capabilities that are understandable, comprehensive, repeatable and policy-relevant. By using various sources of information on AI, including expert evaluations, the project aims at providing policy makers with the knowledge they need to shape future-oriented education- and labour-market policies. The report shows significant improvements in AI’s capabilities in reading since 2016, which reflect **advancements in natural language processing (NLP)** in recent years. AI’s capabilities to solve mathematical tasks have not improved at the same rate. However, experts **predict that increasing investments in AI research and development will lead to significant advancements of AI in both reading and mathematics in the coming years.** The report also demonstrates the potential for AI to outperform large portions of the population in reading and maths. This has important implications for employment and education, as workers are likely to face increasing competition from machines in these skill domains in the future. It also highlights the need to strengthen **the foundation skills of the workforce and prepare it to work together with AI in key domains.** By providing an example of how AI capabilities improve with respect to two key cognitive skills of humans, this study emphasises **the importance of periodically and systematically monitoring the evolution of AI capabilities and comparing them to human skills. This will be useful to policy makers, educators, and researchers who are seeking to understand the implications of technological advancements for the future of work and education.**”

(OCDE, 2023)

3. What are the ethical considerations?

3.1 Guidelines from international institutions

"AI systems are not magic. **Humans design, train and guide AI**, whether they are the people that set AI policies and strategies, the software programmers who build AI systems, those that collect and tag the data used by them, or the individuals who interact with them. This means that **all stakeholders in the AI technology development ecosystem need to understand the key challenges** that require them to contribute to the responsible use of AI. This can include **knowing why and how an AI system was designed, by whom and for what purpose.**" (UNICEF, 2021)

"Regarding the ethical use of AI and data in teaching, learning and assessment, the European Commission recommends four key considerations to be addressed:

- **human action and human control:** capacity for action of education actors (teachers, learners, heads of institutions);
- **fairness:** impartiality, inclusion, non-discrimination and fair distribution of rights and responsibilities;
- **human-centred:** consideration for people, their identity, integrity and dignity;
- **justification of decisions, transparency and explainability:** use of knowledge, facts and data to justify collective decisions in the school environment, through participatory and collaborative decision-making models."

Source: (Directorate General for Education, Youth, Sport and Culture & European Commission, 2022)

The Council of Europe puts forward several principles for the use of AI in education that respects **fundamental rights**:

- *"Right to human dignity.* Teaching, assessment and accreditation should not be delegated to an AI system.
- *Right to autonomy.* Children should have the right to avoid being individually profiled, to avoid dictated learning paths, and to have their development and future life safeguarded.
- *The right to be heard.* Children should have the right not to engage in the use of an AI system without it having an adverse impact on their education.

- *The right not to be discriminated against.* All children should have the opportunity to benefit from the use of technology, not only those from socio-economic groups that can afford it.
- *Right to privacy and data protection.* Children should have the right for their data not to be aggregated and used for commercial purposes without their direct benefit.
- *Right to transparency and explainability.* Children and their parents must be able to understand and challenge any decision made by an AI system."

Source: (Holmes et al., 2022)

3.2 The diversity of ethical issues identified by research

"The ethical and critical issues raised by AI in education are manifold and of diverse origin: use of **big data**, which can have an impact on the **respect of privacy** of students and school staff, production by private companies rather than school authorities, research in computer science or sciences, technology, engineering and mathematics rather than educational sciences, **cognitive bias** due to lack of diversity within design teams."

Source: (Collin & Marceau, 2021)

3.3 The explainability issue

"In the field of education, there are two categories of user who need to take ownership of the digital environment: students and teachers. **It is crucial that teachers take ownership of CBLEs (computer-based learning environments) to promote the wider use of these tools in teaching.** It is therefore necessary to design systems that teachers can adapt to meet their needs. To do so, **teachers must be able to understand the decisions or recommendations made by the AI system concerning their students,** trust the system and, in some way, 'become one' with it so as to explain the combined behaviour of the system as implemented. This calls for an explicit representation of knowledge and decision-making processes. The issue of explainability must be taken into account from the design stage."

(Association Française pour l'Intelligence Artificielle (AFIA), 2022)

4. What are the areas of application?

4.1 Categorisation and taxonomy

Categorisation proposed by the European Commission

| Teaching | Student-focused aid | Teacher-focused aid | Institution-focused aid |
|--|--|--|---|
| <ul style="list-style-type: none"> • Computer-based environment for human learning • Dialogue-based tutoring systems • Language-learning applications | <ul style="list-style-type: none"> • Exploratory learning environments • Automatic formative assessment • AI-powered collaborative learning | <ul style="list-style-type: none"> • Summative assessment of writing, grading • Keyword detection in learner forums • AI teaching assistants • Educational resource recommendation | <ul style="list-style-type: none"> • Exploration of educational data for resource allocation • Diagnosis of learning difficulties • Orientation services |

Source: (European Commission. Directorate-General for Education, Youth, Sport and Culture, 2022)

Further categorisation from the digital theme groups:

- Adaptive and personalised systems (recommendation and selection of content, resources), adaptive assessment and automatic correction
- Chatbots and intelligent teaching systems, e.g. for teaching the investigative approach and scientific approach (Cisel & Baron, 2019)
- Early anticipation, prediction and prevention of school dropout
- Detection and remediation of learning disabilities
- Use of knowledge produced by research using AI (image analysis, annotation, classification, automatic summary/synthesis, processing of big data on textual or image corpora)
- Administrative uses

Source: (DNE-TN2, 2021) (DNE-TN2, 2022) (Lepage Alexandre, 2022) (Romero et al., 2021)

Taxonomy based on the literature review by (Holmes et al., 2022)

Student-focused AIED

| |
|---|
| Intelligent Tutoring Systems (ITS) |
| AI-assisted Apps (e.g., maths, text-to-speech, language learning) |
| AI-assisted Simulations (e.g., games-based learning, VR, AR) |
| AI to Support Learners with Disabilities |
| Automatic Essay Writing (AEW) |
| Chatbots |
| Automatic Formative Assessment (AFA) |
| Learning Network Orchestrators |
| Dialogue-based Tutoring Systems (DBTS) |
| Exploratory Learning Environments (ELE) |
| AI-assisted Lifelong Learning Assistant |

Teacher-focused AIED

| |
|--|
| Plagiarism detection |
| Smart Curation of Learning Materials |
| Classroom Monitoring |
| Automatic Summative Assessment |
| AI Teaching Assistant (including assessment assistant) |
| Classroom Orchestration |

Institution-focused AIED

| |
|---|
| Admissions (e.g., student selection) |
| Course-planning, Scheduling, Timetabling |
| School Security |
| Identifying Dropouts and Students at risk |
| e-Proctoring |

Source: (Holmes et al., 2022) translation by (Bocquet, 2023)

Lastly, (Ouyang & Jiao, 2021) proposes categorisation according to three paradigms:

| | Theoretical underpinning | Implementations | AI techniques | Examples |
|--|--|--|--|--|
| Paradigm 1: AI-directed, learner-as-recipient | Behaviourism <i>[AI represents knowledge of the field and directs learning processes.]</i> | Earlier work of Intelligent Tutoring Systems | AI based on statistical relational techniques | ACT Program Tutor (Anderson et al., 1990); Stat Lady (Shute, 1995) |
| Paradigm 2: AI-supported, learner-as-collaborator | Cognitive, social constructivism <i>[The AI system as an aid, while the learner collaborates to focus on the learning process.]</i> | <ul style="list-style-type: none"> • Dialogue-based Tutoring Systems • Exploratory Learning Environments | Bayesian network, natural language processing, Markov decision trees | An exploratory environment QUE (Metzler & Martincic, 1998) |
| Paradigm 3: AI-empowered, learner-as-leader | Connectivism, Complex adaptive system <i>[AI for boosting human intelligence, the learner is central to the system.]</i> | <ul style="list-style-type: none"> • The human-computer cooperation • Personalized/adaptive learning | The brain-computer interface – Machine learning, deep learning | A real-time MOOC on predictive modelling (Le et al., 2018) |

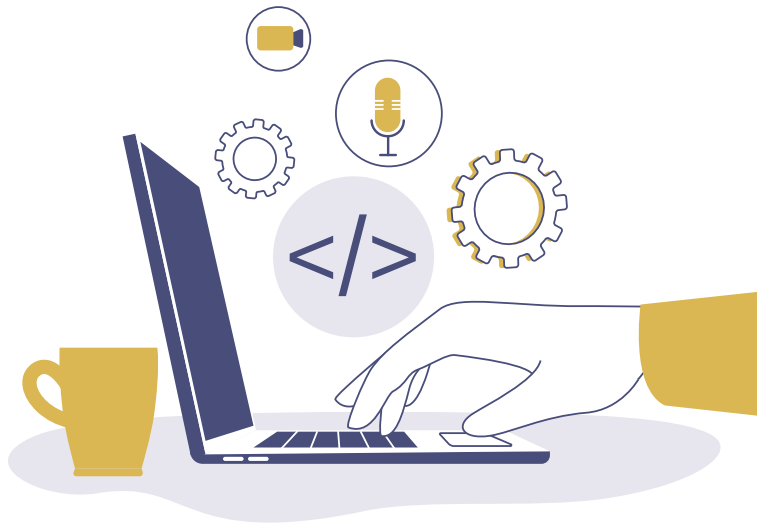
Source: (Ouyang & Jiao, 2021)

4.2 How does it contribute to the learning experience?

“Artificial intelligence (AI) has a large and increasing role in education. One important case is of **personalised teaching systems** which are already well established, with **growing evidence of their effectiveness for improving learning** (VanLehn, 2011; Kulik & Fletcher, 2016; Steenbergen-Hu & Cooper, 2014, 2013; Ma, Adesope, Nesbit, & Liu, 2014; du Boulay, 2016). AI in education (AIED) systems may also make **diverse and sophisticated use of AI to create the interface that is so important for the learning experience**. For example, the interface may use **natural language processing and generation, speech interfaces, avatars, video analysis of the learner** to judge their attention and emotion.”

(Khosravi et al., 2022)

“By using **algorithms, learning can be adaptive**. By analysing the student’s learning traces, such as results from questionnaires or software usage data, the system can modify how it functions to adapt to the user, including in the areas of **content selection and difficulty level**. It is beginning to be possible to analyse learner behaviour using sensors, external ones like a camera, and more intrusive ones, such as a brain-computer interface. The principle of adaptation is at the heart of digital pedagogy, and is most often encountered in a context where **gamification objectives** are also pursued; the learner then enters into an educational game with the machine, sometimes in collaboration with other learners (Giraudon et al., 2020). The KidLearn project (...) offers a learning activity whose multiple variants involving the addition or subtraction of integers or decimals have been designed and implemented by teachers. Variants are organised in the form of a graph of increasing difficulty, respecting the concept of the zone of proximal development (Vygotsky, 1978). This concept is based on the idea that



between an exercise that is too difficult and offputting and an exercise that is too easy and boring, there is an optimal zone that maximises learning progress, measured here by observing the student’s performance throughout the game. It is these elements that are integrated into the algorithm, which will automatically adapt to the learner (Oudeyer *et al.*, 2020).”

(Romero *et al.*, 2023)

Example of adaptive teaching for reading and French language

“Lalilo is developing a teaching assistant for CP, CE1 and CE2 teachers to **help them modify how they teach reading based on each student’s level**, using artificial intelligence technologies. The assistant is an online platform, with a student interface and a teacher interface. It is co-developed with school teachers: the platform is continuously improved, taking into account teacher feedback. (...)”

An adaptive learning algorithm: The goal of Lalilo, to help children learn to read at their own pace, is based on a predetermined learning progression, divided into lessons. Each lesson contains several ‘learning outcomes’, which each contain a wide range of questions. For example, the ‘[a] vowel’ lesson contains two learning outcomes: ‘recognise the grapheme [a]’ and ‘decode the phoneme [a]’. (...) The adaptive learning system allows a child to choose the most suitable exercise at any given time based on their progress and the exercises completed.”

(Gelin, 2022)

[Lalilo is one of the award-winning projects from the Artificial Intelligence Innovation Partnership (P2IA) supported by the French Ministry of National Education and Youth.]

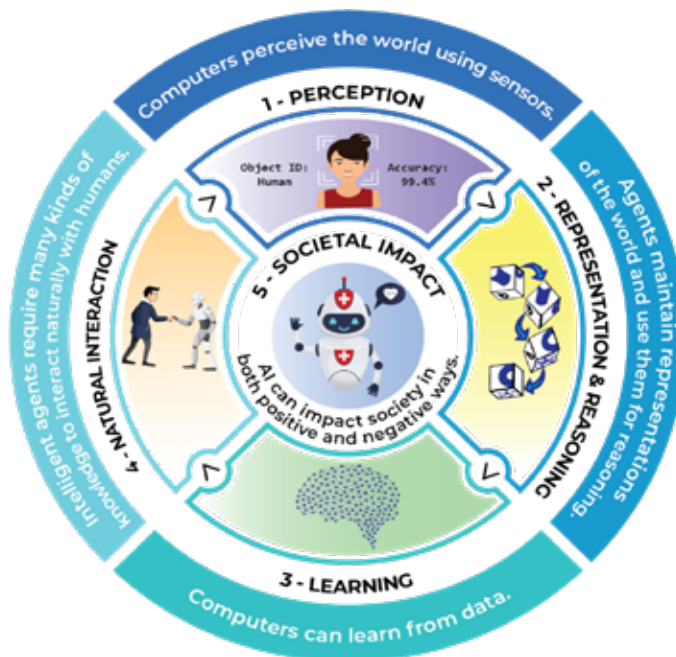
4.3 AI and knowledge production: the example of digital humanities

"...Corpus linguistics and semantics have developed significantly thanks to AI analyses on knowledge representation and machine learning. Likewise, many developments have opened up **opportunities in the area of literary studies**. At the start of the millennium, under the pressure of publishers (not scientists), the term *Humanities Computing*, which referred to research at the fringes of humanities in the 'Anglo-Saxon' sense, **i.e. cultural sciences**, became *Digital Humanities*. This name change meant that humanities were no longer merely instrumented by computers, but were becoming digital. We spoke about a '**computational turning point**' to describe this fundamental evolution. The subsequent change was very successful. In every country, multiple research teams working in the field of classical humanities, whether in history, archaeology or literature, refreshed their approaches in collaboration with computer scientists."

(Ganascia, 2022)

5. What solutions exist for being trained on and teaching AI?

5.1 Opportunities recommended by AI4K12



1. Perception: Computers perceive the world using sensors.

2. Representation & Reasoning: Agents maintain representations of the world and use them for reasoning. Computers construct representations using data structures, and these representations support reasoning algorithms. While AI agents can reason about very complex problems, they do not think the way a human does.

3. Learning: Computers can learn from data thanks to learning algorithms.

4. Natural interaction: Intelligent agents require many kinds of knowledge to interact naturally with humans. Agents must be able to converse in human languages, recognize facial expressions and emotions, but they lack the general reasoning and conversational capabilities of even a child.

5. Societal Impact: AI can impact society in both positive and negative ways. AI technologies are changing the ways we work, travel, communicate, and care for each other. But we must be mindful of the harms that can potentially occur.

Source: (Association for the Advancement of Artificial Intelligence & Computer Science Teachers Association, 2020)

5.2 AI literacy skills framework

Description/learning outcomes

1. Recognizing AI

Distinguish between technological artefacts that use and do not use AI.

2. Understanding intelligence

Critically analyse and discuss features that make an entity 'intelligent'.
Discuss differences between human, animal, and machine intelligence.

3. Interdisciplinarity

Recognize that there are many ways to think about and develop 'intelligent' machines.
Identify a variety of technologies that use AI, including technology spanning cognitive systems, robotics and ML.

4. General vs narrow AI

Distinguish between general and narrow AI.

5. AI strengths and weaknesses

Identify problem types that AI does/does not excel at. Determine when it is appropriate to use AI and when to leverage human skills.

6. Imagine future AI

Imagine possible future applications of AI and consider the effects of such applications on the world.

7. Representations

Understand what a knowledge representation is and describe some examples of knowledge representations.

8. Decision-making

Recognize and describe examples of how computers reason and make decisions.

9. ML steps

Understand the steps involved in machine learning and the practices and challenges that each step entails.

10. Human role in AI

Recognize that humans play an important role in programming, choosing models, and fine-tuning AI systems.

11. Data literacy

Understand basic data literacy concepts.

12. Learning from data

Recognize that computers often learn from data (including one's own data).

13. Critically interpreting data

Understand that data requires interpretation. Describe how the training examples provided in an initial dataset can affect the results of an algorithm.

14. Action and reaction

Understand that some AI systems have the ability to physically act on the world. This action can be directed by higher-level reasoning (e.g. walking along a planned path) or reactive impulses (e.g. jumping backwards to avoid a sensed obstacle).

15. Sensors

Understand what sensors are and that computers perceive the world using sensors. Identify sensors on a variety of devices. Recognize that different sensors support different types of representation and reasoning about the world.

16. Éthique

Identify and describe different perspectives on the key ethical issues surrounding AI: privacy, employment, misinformation, 'singularity',¹¹ decision-making, diversity, bias, transparency and accountability.

17. Programmability

Understand that agents are programmable.

Source: (Long and Magerko, 2020), cited by (UNESCO, 2022)

5.3 AI training for citizens

"AI concerns all citizens, firstly because it is increasingly present in all areas of society and second because **it is necessary to understand it in order to participate in the ethical discussions** it engenders. **It is normal to question the advisability of entrusting algorithms with tasks resulting in crucial decisions**, for example in matters of justice, hiring or other situations with high human consequences. (...) **The MOOC [‘Artificial Intelligence with Intelligence’ from INRIA] adopts a playful and experimental educational approach.** (...) Participants are set hands-on activities: they use neural networks, submit drawings for recognition by AI, are asked to train machine learning models. They are invited to reflect collectively, through discussion forums, on issues raised by the development of AI."

(Romero et al., 2023)

Production of an AI induction module by three chatbots

Response produced by ChatGPT (06/12/23)

Prompt

In the context of school education in France. Recommend, in table format, an induction training module of approximately one hour on artificial intelligence for teachers (primary and secondary education). This module should present the origins of AI, basic definitions, challenges in the field of education, ethical issues and examples of pedagogical application.

Response

"Of course, here is a proposal for an introduction module to artificial intelligence for primary and secondary school teachers in France, presented in table format:

| Time | Content | Objectives |
|-----------|---|--|
| 0-5 min | Introduction | Raise teachers' awareness of the importance of artificial intelligence in the modern world. Briefly present the origins of AI. |
| 5-15 min | Basic definitions | Explain the basic concepts of artificial intelligence such as algorithms, machine learning and neural networks. Use simple terms to ensure basic comprehension. |
| 15-25 min | Challenges in the field of education | Discuss the potential impact of AI on education. Highlight benefits such as personalised learning and challenges such as the need to acquire new skills. |
| 25-35 min | Ethical issues | Examine ethical considerations related to the use of AI in education, with a focus on data privacy, algorithmic discrimination and transparency. Encourage discussion on teachers' ethical responsibility. |
| 35-45 min | Examples of pedagogical application | Present concrete examples of how AI is being used in education, such as intelligent tutorials, automated student performance analysis and learning support tools. Encourage teachers to think about how they could integrate these technologies into their teaching. |
| 45-55 min | Questions and discussions | Invite teachers to ask questions and share their concerns. Facilitate a discussion on how AI can be positively integrated into the educational environment. |
| 55-60 min | Conclusion | Summarise the key points of the module. Provide additional resources to deepen knowledge. Encourage teachers to stay informed about future AI developments in education. |

This module aims to offer a balanced and accessible introduction to artificial intelligence while focusing on aspects relevant to primary and secondary school teachers in France.

Source: <https://chat.openai.com/chat>

Response produced by Llama (07/12/23)

Prompt

In the context of school education in France. Recommend, in table format, an induction training module of approximately one hour on artificial intelligence for teachers (primary and secondary education). This module should present the origins of AI, basic definitions, challenges in the field of education, ethical issues and examples of pedagogical application.

Response

"Here is an example of an approximately one-hour training module for teachers on artificial intelligence:

| Duration | Section | Description |
|----------|--------------------------------------|---|
| 10 min | Introduction | Introduction to the module and objective: introduce teachers to artificial intelligence and its pedagogical applications. |
| 20 min | Origins of AI | Presentation of the origins of artificial intelligence, early concepts and historical developments. |
| 20 min | Basic definitions | Explanations of key terms such as machine learning, deep learning, big data, etc. |
| 15 min | Challenges in the field of education | Discussion of the challenges of artificial intelligence for education, including personalised learning, the effectiveness of teaching methods and accessibility of educational resources. |
| 20 min | Ethical issues | Presentation of ethical considerations related to artificial intelligence, such as privacy, bias, responsibility, etc. |
| 25 min | Examples of pedagogical application | Presentation of concrete examples of the application of artificial intelligence in primary and secondary education, such as the use of chatbots to help students learn, data analysis to improve class management, etc. |
| 10 min | Conclusion | Assessment of the session and recommendation of actions to further explore artificial intelligence in education. |

This module can be adapted to the needs and levels of teachers, and can be supplemented with additional resources such as videos, articles and practical exercises."

Source: <https://poe.com/>

Response produced by Claude (07/12/23)

"Here is a table showing an approximately one-hour introductory training module for teachers on artificial intelligence:

| Duration | Content | Objectives |
|----------|--|---|
| 5 min | Introduction: presentation of the module and its AI awareness objective | Situate the training framework. |
| 10 min | History and basic definitions of AI | Understand the origins and key concepts of AI. |
| 10 min | AI education issues | Identify the potential and challenges of AI in teaching and learning. |
| 10 min | AI ethical issues | Be aware of the ethical issues raised by AI development. |
| 15 min | Concrete examples of pedagogical applications of AI | Illustrate how AI can improve teaching. |
| 10 min | Assessment and outlook | Summarize the key points and the future of teaching and learning. |

This module takes a synthetic but concrete look at the fundamentals of AI and its impact on education, to make teachers aware of the opportunities and challenges posed by this technology. There will also be time for questions and answers."

Source: <https://poe.com/>

5.4 A proposed five-pillar pathway

"Artificial intelligence has not reached maturity. The subject as defined in 1956, studied for 40 years, has achieved spectacular results since 2012 and is still difficult to fully comprehend. It is even more difficult to predict how technologies will evolve, even in the near future. While the development of a complete curriculum is outside the scope of this document, five pillars can be proposed and built upon. (...)

1. Uncertainty and randomness

Data is inconsistent. It does not demonstrate a strict causal relationship.

2. Coding and calculation

Coding and informatics are now part of the school curricula in many countries, in line with recommendations by experts.

3. Data awareness

Data education (data science) will be based on activities where data is collected and viewed, processed and analysed.

4. Critical thinking

Social science can and must contribute to the many ethical issues raised by AI. Critical thinking is an important aspect, although it must always be supported by a proper understanding of how technology works.

5. Post-AI humanism

This is based on the idea that advances in AI are prompting us, as humans, to reconsider some basic truths."

Source: (Higuera & Bocquet, 2020)

5.5 AI training and school dropout prevention

“La **Scientothèque** is a non-profit organisation located within the Experimentarium, the physics museum at the Université Libre de Bruxelles (Belgium). Since it opened, in 2001, its main mission has been to contribute to the reduction of social inequalities. Two realities prompted the establishment of the organisation: on the one hand, the increased risk of school dropouts and the difficulty of accessing higher education for young people from disadvantaged backgrounds (Coslin, 2012) and, on the other hand, the low takeup of scientific studies among girls and women. It has also been shown that it is harder to attract girls to projects of a scientific or technological nature (Blanchard, 2021). (...) A selection of activities was offered to two groups of 10-12-year-olds during the **2020-2021 school year at weekly after-school sessions** at the Institut des Ursulines, a school in Koekelberg, Belgium. The project, funded by the Brussels-Capitale region, aimed to reduce the school dropout rate among pupils from more deprived neighbourhoods. In the sessions, the students were able, through **a fun and collaborative approach**, to learn about the history of machines, explore the basics of biological intelligence, understand the meaning of the word ‘algorithm’, observe and program Thymio robots, get started with programming on the Scratch platform, and discuss the consequences of AI on society.”

(Romero *et al.*, 2023)

5.6 The European “AI4T” project: the “Artificial Intelligence for and by teachers” MOOC

“This course aims to help teachers and the wider educational community gain a general understanding of AI – its uses, technologies and limitations – and to guide them towards an informed use of AI-based tools in the context of education.

The MOOC is organised in four modules:

- Module 1: AI in education
- Module 2: What is meant by AI?
- Module 3: How does AI work?
- Module 4: AI at our service as teachers?

Each module contains texts, videos, activities, quizzes and access to the forum for participants to share their experiences and interact with the teaching team and other participants. The course takes about two or three hours to complete.

Before and after these four course modules, the ‘Overview’ section presents the training objectives and guides participants on how to use the learning platform.

The 'Conclusion' section addresses the ethical aspects of using AI in education. (...)



This course was developed and tested with teachers and school heads as part of the **European AI4T project, bringing together five countries: France, Italy, Ireland, Luxembourg and Slovenia.**

The content was developed together with the five ministries of education involved in the project and in particular the MENJ (Directorate for Digital Education).“ The MOOC is accompanied by an AI4T Open Textbook via the UNESCO RELIA Chair:

<https://chaireunescorelia.univ-nantes.fr/2023/11/16/le-manuel-ouvert-ai4t-en-route-pour-la-deuxieme-edition/>

<https://pressbooks.pub/iapourlesenseignants>

Source: (Inria & AI4T, 2023)



6 The tipping point of generative AI systems and large language models

6.1 The development of generative AI

Presentation

“Before 2014, deep learning models were primarily descriptive, focusing on the summary or patterns and relationships between existing data. These models were designed to explain data and prediction patterns based on the information supplied. However, (Goodfellow *et al.*, 2014) introduced the **Generative Adversarial Network (GAN)**, opening a new era of generative artificial intelligence (GAI). Unlike their descriptive counterparts, generative models, such as GANs, are designed to learn the underlying probability distribution of data. Their main purpose is **to generate new data samples that match as close as possible the patterns observed in the training data**. The development of GANs marked a significant breakaway from traditional deep learning methods, opening up exciting possibilities for generative artificial intelligence. Generative artificial intelligence has since gained huge attention due to its **transformative impact** in various areas of life. It offers seamless solutions to complex problems, enabling the creation of synthetic data, artistic content and realistic images. This paradigm shift in AI technology has profoundly influenced the perception, implementation and use of artificial intelligence, sparking innovation and new application possibilities across industries. (...) The innovative work of Vaswani *et al.* **‘Attention Is All You Need’** [2017] by the Google Brain team introduced a **transformer model** capable of analysing data sets at scale. [This model] was initially developed for natural language processing (NLP), but was then adapted to other areas of machine learning, such as computer vision. (...) The solution offered a new way of handling tasks such as translation, content generation and emotion analysis.”

(Bengesi *et al.*, 2023)

“During the last two years there has been a **plethora of large generative models** such as ChatGPT or Stable Diffusion that have been published. Concretely, these models are able to perform tasks such as being a general question and answering system or automatically creating artistic images that are **revolutionizing several sectors**. Consequently, the implications that these generative models have in the industry and society are enormous, as several job positions maybe transformed. For example, Generative AI is capable of transforming effectively and creatively texts to images, like the DALLE-2 model; text to 3D images, like the Dreamfusion model; images to text, like the Flamingo model; texts to video, like the Phenaki model; texts to audio,

like the AudioLM model; texts to other texts, like ChatGPT; texts to code, like the Codex model; texts to scientific texts, like the Galactica model or even create algorithms like AlphaTensor.”

Generative AI refers to AI technologies that are trained on big data (entire content of Wikipedia, Github, social media, Google images, etc.) and which can generate new content (text, image, audio, video) randomly from prompts entered by users.

Generative AI models (deep neural networks, transformers and other models, such as generative adversarial networks (GANs), which have been increasing in number, especially since 2022) can be broken down according to the following taxonomy:

- Text to image (e.g. DALL-E 2 from OpenAI)
- Text to 3D representation (e.g. Dreamfusion from Google Research)
- Image to text (e.g. Flamingo from Deepmind)
- Text to video (e.g. Phenaki from Google Research)
- Text to audio (e.g. Jukebox from OpenAI)
- Text to text (e.g. ChatGPT from OpenAI)
- Text to code (e.g. Codex from OpenAI)
- Text to scientific formula (e.g. Galactica from Meta AI)
- Other models.

Source: (Gozalo-Brizuela & Garrido-Merchan, 2023)

Foundation models

“A Foundation Model, as named by Stanford University, is a large-scale model based on a deep neural network architecture, trained on enormous quantities of unlabelled data (usually through self-supervised learning). Large Language Models (LLM) are special examples of foundation models that are trained on a corpus of texts. Foundation models open up **fresh perspectives and introduce a new paradigm in language processing, but also in multimodal signal processing** (sound, image, video, etc.). These models that are pre-trained on large corpora can be optimised for a new application using small quantities of additional data specific to that task.”

(CNPEN, 2023)

6.2 Large language models and the new generation of chatbots

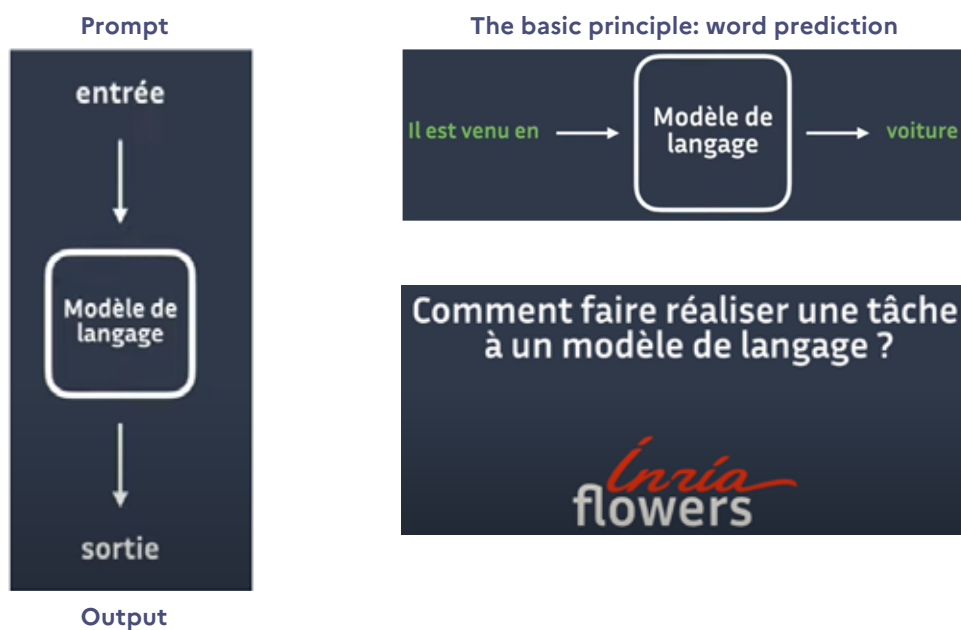
How do large language models work?

The concept of large language models (LLM) is **nothing new**: this type of algorithm is already used on a daily basis in **message assistants**. They work by **predicting the next word** from a prompt entered by the user. The **training phase** allows internal parameters to be adjusted using text data from the web (unverified). A huge amount of data and parameters can improve prediction performance. However, we still can't say that an algorithm "understands" or has a "personality": **responses are generated from statistical regularities**. In addition, these models can generate "**hallucinations**" by inventing false information, which requires extra attention from the user. Currently, the most powerful language models are owned by large companies, which tend to keep any information on their training data to a minimum.

Source: (Inria Flowers, 2023a)

See also (Inria Flowers, 2023b) on prompting: "instructing a language model to perform a certain task (...) by explaining this task to them using phrases in natural language." Two methods are presented: "explain the task in natural language, and give examples of how to complete the task."

Operation of a large language model and prompting



Source: (Inria Flowers, 2023)

The “turning point of ChatGPT”: origins and technology foundation

In this category of generative AI models, the application ChatGPT, which OpenAI launched publicly on 30 November 2022, represents a **new generation of chatbots**, “agents that converse with a user to carry out research, find an answer to a question or even practise a given discipline” (CSEN & DNE-TN2, 2022). Due to the performance shown, it is clearly a **“qualitative leap forward”** in automatic language processing (ALP) - or **Natural Language Processing (NLP)** technologies (Boisson & DNE-TN2, 2023).

This **“chatbot prototype (...)**, a language model refined using supervised learning and reinforcement learning techniques” (“ChatGPT”, 2023), based on generative artificial intelligence, capable of generating on-demand texts or computer code from questions in natural language, is a **proprietary application** developed from an optimisation of the GPT 3.5 algorithmic model.

“GPT-3 is larger than GPT-2, with 175 billion parameters. To get a more accurate picture of the size of this algorithm, Wikipedia’s data, which includes around 6 million articles, represents only 0.06% of the data trained by GPT-3.”

(Boa Sorte *et al.*, 2021)

The technologies underpinning ChatGPT can be summarised as follows:

- it is a **“large language model... a type of AI that uses deep learning (a form of machine learning) to process and generate natural language texts (...)** [These models are] trained on **massive amounts of text data, allowing them to learn the nuances and complexities of human language”** (Source: Susnjak, 2022);
- its name is a contraction of “chat” and “GPT” [Generative Pre-trained Transformer], which is a “predictive language translation model” (transformation of data represented in one form into data represented in another form);
- its **transformer-based** algorithm (Vaswani *et al.*, 2017) is a neural network architecture model that uses the **self-attention mechanism to establish relationships between words and determine the meaning of a word according to the context**, focusing on the most relevant words, thus allowing a credible approach to natural language;
- it works according **to a word prediction model**, based on the prompts given by the user in natural language, combining in its training phase learning by reinforcement and supervised learning from human feedback;
- it is therefore important to remember that, like any AI technology, texts are generated and **reinforced by human agency and efforts** (from the people who design the algorithm to those who identify or verify data or interact with the program).

The technology foundation of ChatGPT is the result of **“a combination of two models: a language model** (its references ended in 2021), GPT, a giant text prediction model and encyclopaedic model, and a conversational model (refined based on user feedback), InstructGPT, an ‘improved’ version of GPT-3 from the labelling of generated texts.”

It is also the result of **“a series of strata from linguistic and informatics research:**

- **semantic statistics**, with the concept of ‘popup window’ (the meaning comes from the relative position of the words with each other);
- **word vectorisation** (word embedding) allowing the characterisation of a word within contextual spaces in relation to neighbouring words;

- **neural networks** to recognise precise syntactic and semantic relationships between words and manage language transfers (responses to user prompts can thus be based on a multilingual semantic space);
- **transformer models** (2017) based on the attention mechanism and modelling syntactic formulas;
- the development of **large language models**, starting in 2018 with BERT (Google), then GPT-2 and 3 (OpenAI)."

Source: (Langlais, 2023)

On 14 March 2023, OpenAI announced the upgrade of its model with the launch of **GTP-4** (fee-based service or available in the new version of the Bing search engine): "a large-scale **multimodal model** that can accept image and text inputs and produce text outputs. Although less capable than humans in many real-world scenarios, GPT-4 exhibits human-level performance across a variety of professional and academic criteria, including passing a simulated bar examination with a score achieved by around 10% of top candidates." (OpenAI, 2023)

Other large language models

Other large language models include, in particular, in open source:

- **Bert** (Bidirectional Encoder Representations from Transformers), developed by Google in 2018 (Devlin *et al.*, 2019);
- **Bloom**, initiated by the French-American start-up Hugging Face, trained on the "Jean Zay" machine, "one of the most powerful supercomputers in Europe" (CNRS, 2022);
- **LLaMA** from Meta AI (Touvron *et al.*, 2023).

For more details, see also (Pôle d'expertise de la régulation numérique, 2023).

Examples of state-of-the-art large language models and associated openness policies

| | Year | No. maximum parameters (in billions) | Public architecture | Pretrained open model | Open data | Conversational (RLHF) | Accessible to users (UI or API) |
|---|---------------|--------------------------------------|---------------------|-----------------------|---|-----------------------|---------------------------------|
| BigScience BLOOM | 2022 | 175 | ✓ | ✓ | ✓ | ✗ | ✓ |
| Google GLaM/PaLM | 2021/ 2022 | 1200/540 | ✓ | ✗ | ✗ | ✗ | ✓ Fee-based API |
| Google LaMDA/Bard | 2022 | 137/? | ✓/✗ | ✗ | ✗ | ✓ | ✓ UK/US |
| Meta OPT | 2022 | 175 | ✓ | ✓ | ✓ | ✗ | ✗ |
| Meta BlenderBot3 | 2022 | 175 | ✓ | ✓ | ✓ | ✓ | ✓ US |
| Meta LLaMA | 2023 | 65 | ✓ | ✓ | ✓ | ✗ | ✗ |
| OpenAI GPT-3 | 2020 | 175 | ✓ | ✗ | ✓ 1 ^{re} version seulement | ✗ | ✓ Fee-based option |
| OpenAI GPT-3.5 (InstructGPT /ChatGPT) | 2022 | 175/? | ✓/✗ | ✗ | ✗ | ✓ | ✓ Fee-based option |
| OpenAI GPT-4 | 2023 | ? | ✗ | ✗ | ✗ | ✓ | ✓ Fee-based API and UI |

Source: (Pôle d'expertise de la régulation numérique, 2023)

In December 2023, Google presented **Gemini** (Generalised Multimodal Intelligence Network), its new multimodal artificial intelligence system (Hsiao, 2023).

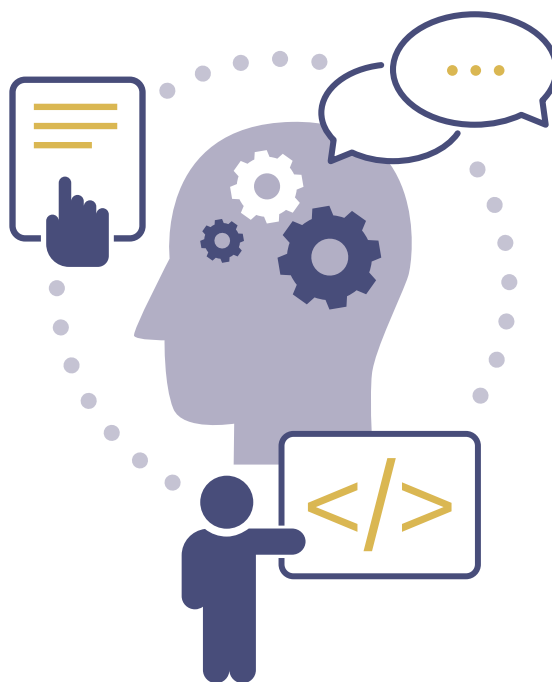
Reactions and questions

Since the launch of the ChatGPT app, there has been a lot of **reactions and debate** in the media and in various professional environments, including research and education.

“These multilingual, multitasking, interactive, generative AIs raise **many issues**: the data chosen to train them, the distribution of languages in a multilingual system, the optimisation parameters of the systems, the ownership of the content generated, etc. In addition, the generative power of AIs is often amplified by **filters that allow the censoring of certain topics** and logical deduction modules in order to verify the veracity of responses. A small group of people (engineers, transcribers, evaluators) have created this type of system used by millions of people. These widely used artificial intelligence systems therefore pose **considerable ethical issues**, including the transformation of the notion of information production, relationship with the truth and the massive risks associated with disinformation and manipulation.”

(Devillers, 2023)

Reactions were a mix of **fascination**, faced with the credible and plausible nature of the “conversations” produced and the relevance – at least apparent – of certain answers, **concern** about possible misuse (cheating on school work, exam fraud, doubt about the authorship of written texts produced in a personal or professional context) and **uncertainties about technologies and the data** mobilised and the reliability of the responses produced.



What are the educational opportunities?

This type of application offers several educational opportunities that are currently being actively explored and shared by the teacher community:

- **pedagogical engineering aid** (design of teaching and learning materials, see proposed matrix below);
- **creativity aid** by proposing, based on the prompts and questions entered, an initial draft or code;
- **assistance with learning foreign languages** through the conversational followup function;
- **assistance** with initiating or optimising a writing-based task;
- a **new challenge** for media and information literacy (MIL) and a framework for teaching critical thinking, which requires teachers and their students to systematically check the content and relevance of the answers provided.

These opportunities have already been the subject of several publications, among which (Kasneci *et al.*, 2023) which highlights the following: personalised learning (analysis of work and suggestion of adapted teaching materials), course planning (assistance with the creation of inclusive plans and activities), and language learning (summaries and translations).

In addition, several **teaching guides** set out ways for the teacher to use them. One guide produced by (Herft, 2023), an Australian teacher, and translated into French, proposes a series of examples of instructions to enter to make it a teaching and learning assistant. e.g. "Create a course plan for a class ... that covers the concept of ... and includes a variety of activities and assessments."

Prompting and pedagogical engineering: example of a matrix for an instruction sheet

| Information | Educational situation | Complete field |
|---|--|----------------|
| Identity Context Discipline/learning area | I am a teacher of [...] in [type of institution] | |
| Framework of action Constraints Difficulties | The school or students has/have [profile], have to deal with [difficulty] | |
| Outcomes Objectives Target audience/level Actions targeted | I teach students of [level, age] – I want to help them improve in or teach them [subject, knowledge, skill] | |
| Question Request | I want to build [sequence], produce [resource, learning situation] for/with my students. | |
| Style Format Keywords Steps | I would like the answer in the form of a [list/table...], to produce a visual, a diagram, in a language adapted to [level], using the following keywords: ... <i>Complete with an example of the expected result if necessary.</i> <i>If necessary, detail the steps to follow to improve the result (chain-of-thought prompting).</i> | |

(Allouche, 2023b)

Fonctions de ChatGPT pour soutenir les enseignants dans leurs pratiques

| Aspect | Function | Representative Quotes |
|----------------------|---------------------------------|--|
| Teaching preparation | Generating course materials | "we asked ChatGPT to create one of the dialogues above in a format that can be imported into DialogFlow. ChatGPT was able to generate it successfully" |
| | Providing suggestions | "I told ChatGPT that the learner was with dyslexia, and eventually ChatGPT recommended specific learning materials for the learner" |
| | Performing language translation | "ChatGPT can be used to translate educational materials into different languages" |
| Assessment | Generating assessment tasks | "Another use of ChatGPT is its ability to generate exercises, quizzes, and scenarios which can be used in the classroom to aid in practice and assessment" |
| | Evaluating student performance | "ChatGPT can be trained to grade student essays, providing teachers with more time to focus on other aspects of teaching" |

Source: (Lo, 2023)

The perspectives of large language models for people with disabilities

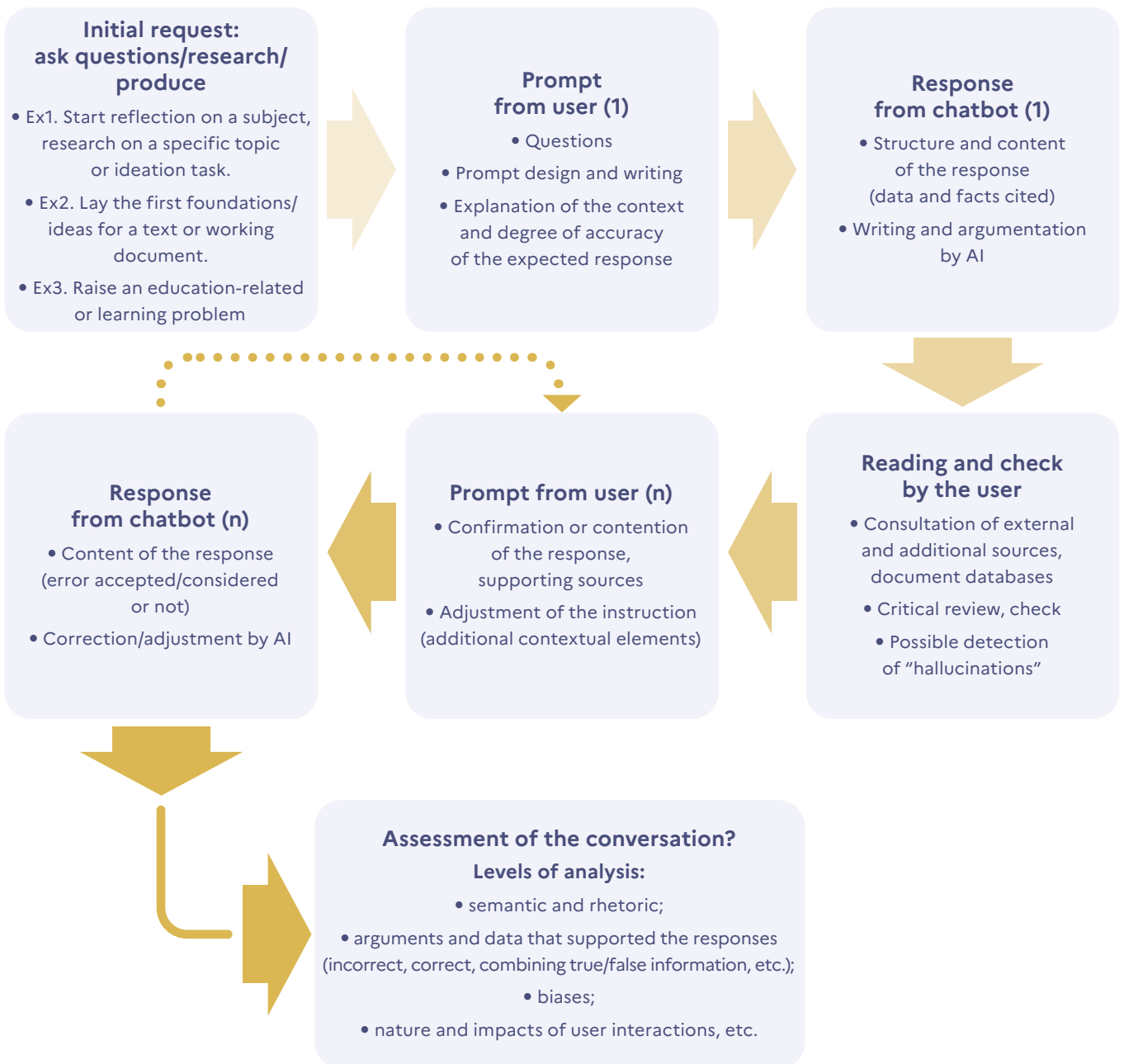
"To empower learners with disabilities, large language models can be used in combination with speech-to-text or text-to-speech solutions to help people with visual impairment. In combination with the previously mentioned group and remote tutoring opportunities, language models can be used **to develop inclusive learning strategies** with adequate support in tasks such as adaptive writing, translating, and highlighting of important content in various formats. However, it is important to note that the use of large language models **should be accompanied by the help of professionals** such as speech therapists, educators, and other specialists that can adapt the technology to the specific needs of the learners disabilities."

(Kasneci *et al.*, 2023)

What limits and areas for consideration (ethical, educational and scientific dimensions)?

Several tests were performed to evaluate the performance and limitations of ChatGPT (GPT 3.5) (Allouche, 2023a). The diagram below proposes a model of a critical user journey, integrating prompts that can be simple or complex, ranging from a simple request for information to problem solving.

Modelling of a critical user journey with a chatbot



(Allouche, 2023a) updated

Following these tests, several limitations and points of caution may be identified (bearing in mind that, under **GDPR**, the application cannot be used in a school setting, only individual uses are permitted under the private contract system):

- for the time being, the responses provided are **not based on real-time data** via a web connection (subject to verification, this is data gathered from the web, prior to 2022);
- they contain **biases** due to the nature of the data processed, even if filters are applied to limit the most obvious biases;
- **factual errors** and possible “hallucinations” require in almost every instance verification of the content by the user, which can be both a constraint and an asset in an educational context for developing critical thinking and creating a healthy distance with the content generated;
- the risk of producing and disseminating truncated, incorrect or unverifiable data or information;
- this kind of technology, owing to its performance and high volume of data processed (without transparency on the nature of said data), inevitably raises the questions of referencing (can we refer to the answers produced when they are not systematically sourced or reproducible?), the authorship of the content (who does it belong to, can the authorship of a work be attributed to an AI engine?) (Stokel-Walker, 2023) or **copyright** (do the responses generated from this big data respect copyright when applicable? e.g. on a scientific or literary work that is not copyright-free);
- the **absence of sources** – or the potentially incorrect nature of those cited – constitutes a major problem for use in an academic context (education, research, documentation), unless the user expressly requests this, and provided that the user checks the references cited one by one.

In response to this problem, other complementary or competing applications, such as Bing Chat (Microsoft’s search engine that now integrates GPT-4) or Perplexity (<https://www.perplexity.ai/>), can provide sources (but without guaranteeing their relevance or completeness).

The **environmental impact** of this type of model, which is not specific to ChatGPT, should also be noted:

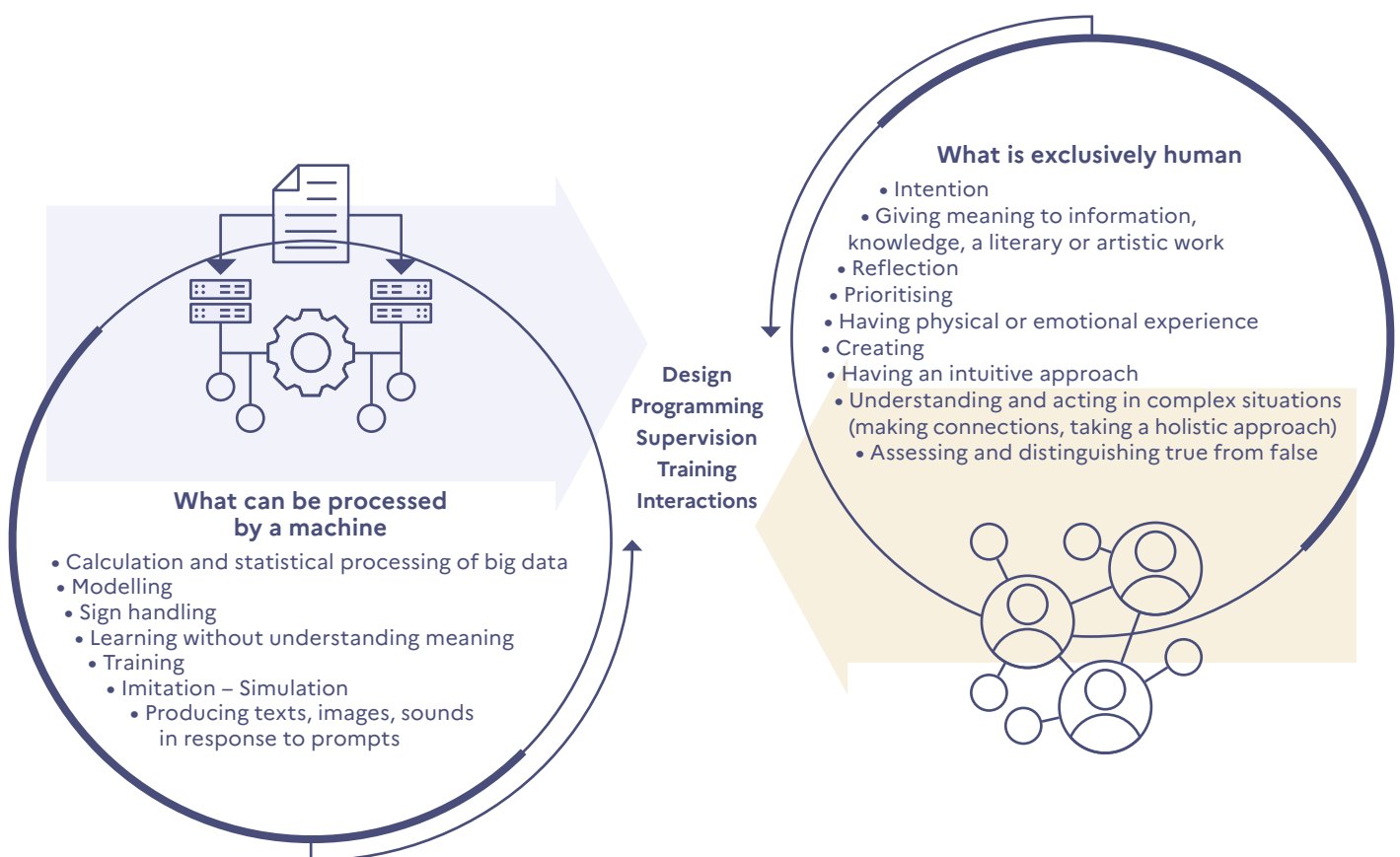
“Training OpenAI’s GPT-3 has been estimated to have required as much energy as driving a car to the moon and back, thus generating an equivalent of 85,000 kg of CO2 emissions (Quach, 2020). Energy consumption is now widely understood to be a major challenge for data-driven AI (Strubell *et al.*, 2019).”

(Holmes & Tuomi, 2022)

In summary, the main opportunities to explore are primarily assistance (see examples pp. 33-35), inspiration, and complementing or furthering knowledge. Indeed, such a technological leap is inviting researchers and teachers in particular to innovate, collaborate, inform and **educate themselves, remotely and with a critical mindset**, to support students in their learning and training.

As the diagram below suggests, one of the challenges to overcome is the **distinction to be made with students between what is exclusively human** – giving meaning to information, knowledge, a literary or artistic work – **and what can reliably be entrusted to AI**: processing and computing big data, beyond human capability, for the purposes of prediction, probability or optimisation, but without intent or comprehension of what is produced.

What may or may not be entrusted to AI



(DNE-TN2, 2023)

Références

Allouche, E. (2023a, janvier 5). Sens et finalités du numérique en éducation – Hors-série : Tests et simulations d'« entretien » avec ChatGPT (Open AI) [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/7635>

Allouche, E. (2023b, octobre 2). IA génératives et ingénierie pédagogique : Le *prompting*, pistes de travail et applications [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/9934>

Association for the Advancement of Artificial Intelligence, & Computer Science Teachers Association. (2020). AI4K12. AI4K12. <https://ai4k12.org/>

Association Française pour l'Intelligence Artificielle (AFIA). (2022, janvier 6). Journée Enseignement et Formation en IA : « IA pour l'enseignement ». AfIA. <https://afia.asso.fr/efia-2022/>

Bengesi, S., El-Sayed, H., Sarker, M. K., Houkpati, Y., Irungu, J., & Oladunni, T. (2023). *Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers* (arXiv:2311.10242). arXiv. <https://doi.org/10.48550/arXiv.2311.10242>

Boa Sorte, P., Farias, M. A. de F., dos Santos, A. E., Santos, J. do C. A., & Dias, J. S. dos S. R. (2021). Artificial intelligence in academic writing: What is in store with the GPT-3 algorithm? *Revista EntreLínguas*, 7(00). <https://periodicos.fclar.unesp.br/entrelinguas/article/download/15352/11410/51225>

Bocquet, F. (2023, janvier 23). État de l'art et de la pratique de l'intelligence artificielle dans l'éducation (Holmes & Tuomi, 2022) [Traduction] [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/8350>

Boisson, J., & DNE-TN2. (2023, janvier 12). Intelligence artificielle, traitement automatique des langues et agents conversationnels [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/8235>

Cardon, D., Cointet, J.-P., & Mazieres, A. (2018). La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, 5(211). <https://doi.org/10.3917/res.211.0173>

Cisel, M., & Baron, G.-L. (2019). Vers des intelligences artificielles pour l'enseignement de la démarche d'investigation. *Educational Journal of the University of Patras UNESCO Chair*, 0, Article 0. <https://doi.org/10.26220/une.2990>

CNPEN. (2023, juin 30). Avis 7 du CNPEN *Systèmes d'intelligence artificielle générative : Enjeux d'éthique*. Comité Consultatif National d'Éthique. <https://www.ccne-ethique.fr/fr/publications/avis-7-du-cnpén-systemes-dintelligence-artificielle-generative-enjeux-dethique>

CNRS. (2022, juillet 12). *Livraison du plus grand modèle de langue multilingue « open science » jamais entraîné*. CNRS. <https://www.cnrs.fr/fr/livraison-du-plus-grand-modele-de-langue-multilingue-open-science-jamais-entraîne>

Collin, S., & Marceau, E. (2021). L'intelligence artificielle en éducation : Enjeux de justice. *Formation et profession*, 29(2), 1. <https://doi.org/10.18162/fp.2021.a230>

Conseil de l'Europe. (2023, décembre 9). Législation sur l'intelligence artificielle : Le Conseil et le Parlement parviennent à un accord sur les premières règles au monde en matière d'IA. Conseil européen - Conseil de l'Union européenne. <https://www.consilium.europa.eu/fr/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>

Cour des comptes. (2023). *La stratégie nationale de recherche en intelligence artificielle*. Cour des comptes. <https://www.ccomptes.fr/fr/publications/la-strategie-nationale-de-recherche-en-intelligence-artificielle>

CSEN, & DNE-TN2. (2022, novembre 7). Synthèse et recommandations du groupe de travail « Intelligence artificielle et numérique pour l'éducation » du CSEN [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/6707>

De la Higuera, C., & Bocquet, F. (2020, juillet 9). L'éducation, la formation des enseignants et l'apprentissage de l'intelligence artificielle : Un aperçu des questions clés [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/1973>

Devlillers, L. (2023, février 20). ChatGPT et « intelligences » artificielles : Comment déceler le vrai du faux. *The Conversation*. <http://theconversation.com/chatgpt-et-intelligences-artificielles-comment-deceler-le-vrai-du-faux-200181>

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv. <https://doi.org/10.48550/arXiv.1810.04805>

Directorate-General for Education, Youth, Sport and Culture, Y., & European Commission. (2022). *Lignes directrices pour les enseignants et les éducateurs en matière de lutte contre la désinformation et de promotion de l'habileté numérique par l'éducation et la formation*. Publications Office of the European Union. <https://op.europa.eu/en/publication-detail/-/publication/a224c235-4843-11ed-92ed-01aa75ed71a1/language-fr>

DNE-TN2. (2021, juin 30). Le deep learning, un tournant épistémologique pour les SHS ? (Huma-Num LAB, 2020) [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/3203>

DNE-TN2. (2022, janvier 20). Les enjeux éducatifs à l'ère de l'Intelligence Artificielle : GTnum #Scol_IA – Maison de l'Intelligence artificielle [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/3898>

DNE-TN2. (2023a, mars 9). Intelligence artificielle et éducation ouverte : Retour sur la journée du GTnum #IA_EO (26/01/23) [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/8605>

DNE-TN2. (2023b, juillet 10). Enseigner et apprendre à l'ère de l'intelligence artificielle : Portfolio du GTnum #Scol_IA [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/9593>

DNE-TN2. (2023c, septembre 6). Intelligence artificielle et éducation ouverte : Portfolio du GTnum #IA_EO [Billet]. *Éducation, numérique et recherche*. <https://edunumrech.hypotheses.org/9781>

- European Commission. Directorate General for Education, Youth, Sport and Culture. (2022). *Lignes directrices éthiques sur l'utilisation de l'intelligence artificielle (IA) et des données dans l'enseignement et l'apprentissage à l'intention des éducateurs*. Publications Office. <https://data.europa.eu/doi/10.2766/153756>
- Ganascia, J.-G. (2022, juin 10). *Intelligence artificielle et épistémologie. Allers-retours indispensables*. JECIS. <https://hal.science/hal-03760357>
- Gelin, L. (2022). *Reconnaissance automatique de la parole d'enfants apprenants-tes lecteurs-trices en salle de classe : Modelisation acoustique de phones* [Thèse, Toulouse 3]. <https://hal.science/tel-03715653v1>
- Gouvernement - Direction générale des Entreprises. (2023). *Éduquer à l'Intelligence artificielle. Stratégie nationale pour l'intelligence artificielle*. <https://www.entreprises.gouv.fr/fr/numerique/enjeux/developper-ecosysteme-de-talents>
- Gozalo-Brizuela, R., & Garrido-Merchan, E. C. (2023). *ChatGPT is not all you need. A State of the Art Review of large Generative AI models*. <https://doi.org/10.48550/ARXIV.2301.04655>
- Grand lexique français de l'Intelligence artificielle*. (2021). DataFranca Wiki. <https://datafranca.org/wiki/Accueil>
- Herft, A. (2023). *Guide de l'enseignant—L'usage de ChatGPT « ce qui marche le mieux »* (A. Gagné, Trad.). <https://labua.univ-angers.fr/wp-content/uploads/2023/01/ChatGPT-Guide-de-lenseignant-FR.pdf>
- Holmes, W., Persson, J., Chounta, I.-A., Wasson, B., & Dimitrova, V. (2022). *Artificial intelligence and education—A critical view through the lens of human rights, democracy and the rule of law*. Council of Europe. <https://rm.coe.int/artificial-intelligence-and-education-a-critical-view-through-the-lens/1680a886bd>
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, n/a(n/a). <https://doi.org/10.1111/ejed.12533>
- Hsiao, S. (2023, décembre 6). *Bard gets its biggest upgrade yet with Gemini*. Google. <https://blog.google/products/bard/google-bard-try-gemini-ai/>
- Inria, & AI4T. (2023). *Intelligence artificielle pour et par les enseignants*. FUN MOOC. <http://www.fun-mooc.fr/fr/cours/intelligence-artificielle-pour-et-par-les-enseignants-ai4t/>
- Inria, & Class'Code. (2020). *L'Intelligence Artificielle... avec intelligence ! (MOOC)*. FUN-MOOC. <https://www.fun-mooc.fr/fr/cours/lintelligence-artificielle-avec-intelligence/>
- Inria Flowers (Réalisateur). (2023a, mars 8). *Comment fonctionne ChatGPT ? Un tour d'horizon en moins de 5mn*. <https://www.youtube.com/watch?v=K8gOvC8gvB4>
- Inria Flowers (Réalisateur). (2023b, mars 13). *Le prompting, ou l'art de se faire comprendre par ChatGPT : Explication en 5mn*. https://www.youtube.com/watch?v=8lQ9i_QoA3A
- Kasneji, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneji, G. (2023). *ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education*. <https://doi.org/10.35542/osf.io/5er8f>

- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Gasevic, D., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Tsai, Y.-S. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence*, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
- Langlais, P.-C. (2023, février 7). ChatGPT : Comment ça marche ? [Billet]. *Sciences communes*. <https://scoms.hypotheses.org/1059>
- Le Blanc, B. (2014). Alan Turing : Les machines à calculer et l'intelligence. *Hermès, La Revue*, 68(1), 123-126. <https://doi.org/10.3917/herm.068.0123>
- Le Cun, Y., Brizard, C., & Brizard, C. (2019). *Quand la machine apprend : La révolution des neurones artificiels et de l'apprentissage profond*. Odile Jacob.
- Lepage Alexandre (Réalisateur). (2022, mars 3). *L'intelligence artificielle en éducation*. <https://www.youtube.com/watch?v=ZAtu0xllcls>
- Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Education Sciences*, 13(4), Article 4. <https://doi.org/10.3390/educsci13040410>
- Lorenz, P., Perset, K., & Berryhill, J. (2023). *Initial policy considerations for generative artificial intelligence*. OCDE. <https://doi.org/10.1787/fae2d1e6-en>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO Bibliothèque Numérique. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- OCDE. (2023, mars 28). *Is Education Losing the Race with Technology?: AI's Progress in Maths and Reading*. OCDE. <https://www.oecd.org/education/is-education-losing-the-race-with-technology-73105f99-en.htm>
- OpenAI. (2023). *GPT-4 Technical Report*. OpenAI. <https://cdn.openai.com/papers/gpt-4.pdf>
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education : The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020. <https://doi.org/10.1016/j.caeai.2021.100020>
- Parlement européen. (2023, décembre 9). *Artificial Intelligence Act: Deal on comprehensive rules for trustworthy AI*. News European Parliament. <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>
- Pôle d'expertise de la régulation numérique. (2023). ChatGPT ou la percée des modèles d'IA conversationnels. *Gouvernement français*, 6. https://www.peren.gouv.fr/rapports/2023-04-06_Eclairage%20sur_CHATGPT_FR.pdf
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 22. <https://doi.org/10.1186/s41039-017-0062-8>
- Romero, M., Aloui, H., Heiser, L., Galindo, L., & Lepage, A. (2021). *Un bref parcours sur les ressources, pratiques et acteurs en IA et éducation*. <https://hal.archives-ouvertes.fr/hal-03190014>

Romero, M., Heiser, L., Lepage, A., Gagnebien, A., Bonjour, A., Lagarrigue, A., Palaude, A., Boulord, C., Gagneur, C.-A., Mercier, C., Caucheteux, C., Guidoni-Stoltz, D., Tressols, F., Alexandre, F., Céci, J.-F., Metral, J.-F., Camponovo, J., Henry, J., Fouché, L., ... Borgne, Y.-A. L. (2023). *Enseigner et apprendre à l'ère de l'intelligence artificielle : Livre blanc du G7num #Scol_IA: Vol. Livre blanc*. <https://hal.science/hal-04013223>

Singh, R., & Gill, S. S. (2023). Edge AI : A survey. *Internet of Things and Cyber-Physical Systems*, 3, 71-92. <https://doi.org/10.1016/j.iotcps.2023.02.004>

Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: Many scientists disapprove. *Nature*. <https://doi.org/10.1038/d41586-023-00107-z>

Susnjak, T. (2022). *ChatGPT: The End of Online Exam Integrity?* (arXiv:2212.09292). arXiv. <https://doi.org/10.48550/arXiv.2212.09292>

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA : Open and Efficient Foundation Language Models* (arXiv:2302.13971). arXiv. <https://doi.org/10.48550/arXiv.2302.13971>

UNESCO. (2019). *Consensus de Beijing sur l'intelligence artificielle et l'éducation*. UNESCO Bibliothèque Numérique. <https://unesdoc.unesco.org/ark:/48223/pf0000368303>

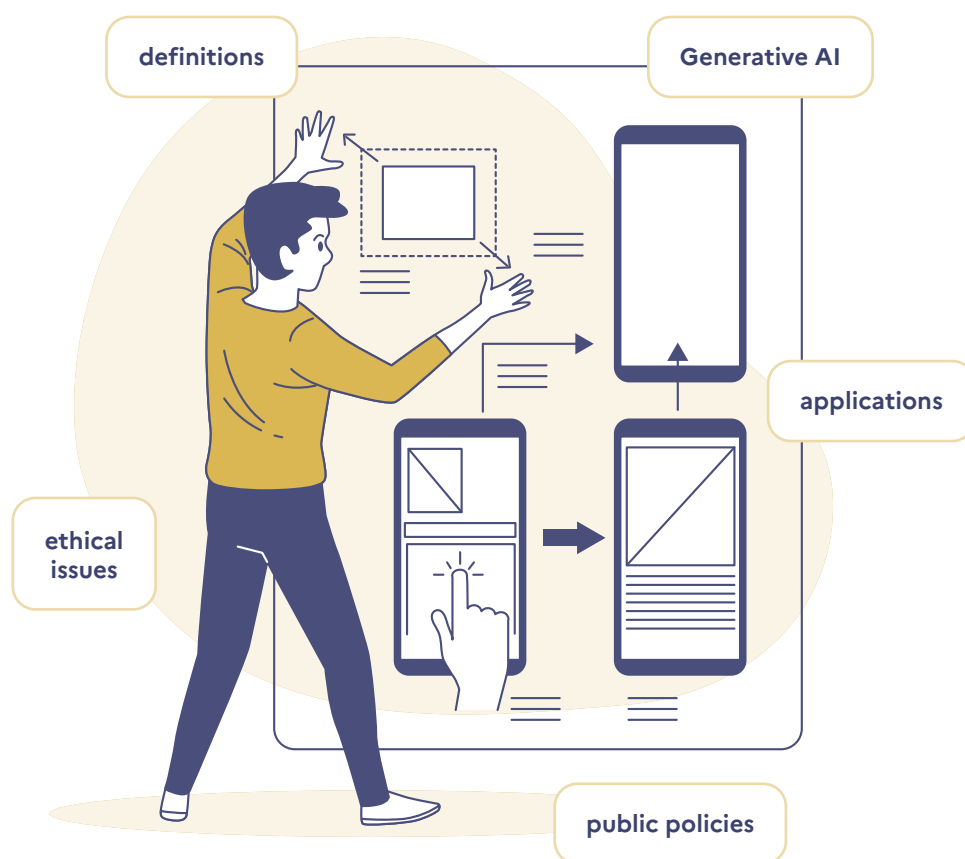
UNESCO. (2021). *IA et éducation : Guide pour les décideurs politiques*. UNESCO Bibliothèque Numérique. <https://unesdoc.unesco.org/ark:/48223/pf0000380006>

UNESCO. (2022). *K-12 AI curricula: A mapping of government-endorsed AI curricula*. UNESCO Bibliothèque Numérique. <https://unesdoc.unesco.org/ark:/48223/pf0000380602>

UNICEF. (2021, novembre). *Orientations stratégiques sur l'IA destinée aux enfants 2.0*. UNICEF. <https://www.unicef.org/globalinsight/fr/rapports/orientations-strat%C3%A9giques-sur-lia-destin%C3%A9e-aux-enfants>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. <https://doi.org/10.48550/ARXIV.1706.03762>

Article reference: DNE-TN2 (2024). Intelligence artificielle et éducation : Apports de la recherche et enjeux pour les politiques publiques (édition 2024) [Billet].
Éducation, numérique et recherche.
<https://edunumrech.hypotheses.org/10764>



Carnet Hypothèses

Education, digital technology and research





<https://edunumrech.hypotheses.org/>

Contact



Direction du numérique pour l'éducation
(Ministère de l'Éducation nationale et de la Jeunesse)
Bureau du soutien à l'innovation numérique
et à la recherche appliquée (DNE TN2)

@Edu_Num

 Licence Ouverte 2.0
 Open Licence 2.0