Consultative Meeting

HOW CONFIDENT?...

Artificial Intelligence, Young People and Youth Work

A consultative meeting about developments in Artificial Intelligence and their impact on young people's human rights, youth work and youth policy in the context of Council of Europe standards and approaches.

22-23 January 2025
European Youth Centre, Strasbourg, France

GLOSSARY OF TERMS

Prepared by Ron Salaj

Defining AI

AI can be broadly defined as a branch of computer science dedicated to creating systems capable of performing tasks that typically require human intelligence. These tasks encompass a wide range of capabilities, from simple algorithms that automate repetitive actions, like sorting emails, to complex systems that can learn, adapt, and make decisions or recommendations.

Over the years, many definitions of AI have been proposed, reflecting its evolving nature and the diversity of its applications. Among these, the Organisation for Economic Co-operation and Development (OECD) offers a widely adopted definition that is now the standard for tthe Council of Europe (CoE) and the European Union (EU):

> *“An Artificial Intelligence system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.”*

This definition of AI may seem abstract at first sight. However, to make sense of it, we shall break it down into smaller pieces revealing how its building blocks are connected and how they relate in everyday life:

- **“A Machine-Based System”:** AI systems rely on computing power to analyse data and perform tasks. These systems use hardware (like processors or sensors) and software (algorithm) to carry out their functions. For example, a smartphone with a voice assistant (e.g. Siri or Google Assistant) is a machine-based system designed to interpret and respond to human speech.
- **“For explicit or implicit objectives”:** AI systems are created with a purpose in mind. Some objectives are explicit, meaning they are clearly defined, like optimising traffic lights to reduce congestion. Others are implicit, where the system learns objectives through patterns in the data, such as identifying user preferences for personalised content. For example, a fitness app explicitly tracks your steps, while implicitly suggesting workouts based on your activity patterns.
- **"Infers, from the input it receives":** AI systems process input data to make sense of the world. This involves analysing patterns, relationships, or trends in the data to make predictions or generate responses. Inputs can include text, images, numbers, or even voice commands. For example, a facial recognition system infers identities from images by analysing patterns in a person’s facial features.
- **"How to generate outputs":** The output is the result or action produced by the AI system based on the processed inputs. Outputs can include predictions, like weather forecasts; content, like auto-generated captions for videos; recommendations, like movie suggestions; or decisions, like approving a loan application. For example, a recommendation system like Netflix produces outputs by suggesting movies or shows you’re likely to enjoy.
- **"That can influence physical or virtual environments":** AI systems are impactful because their outputs can change the way systems or people behave. In physical environments, this might involve robots or self-driving cars. In virtual environments, it could mean influencing what

content you see online or how information is ranked in search engines. For example, a virtual assistant adjusting your smart home's temperature impacts your physical environment, while an AI-based social media algorithm curating your feed influences your virtual experience.

- **"Levels of autonomy and adaptiveness after deployment":** AI systems vary in how much independence they have once deployed. Some require continuous human supervision (e.g., medical diagnostics tools), while others are adaptive, meaning they can learn and improve on their own over time (e.g., language models like Google Translate, Gemini, or Chatgpt).

Illustrative Example: The Smart Home Assistant

Imagine a smart home assistant—let's call it Echo—controlling your home environment. Let's see how it incorporates all the building blocks of AI's definition:

- A Machine-Based System: The assistant uses hardware (microphones, speakers, and sensors) and software (natural language processing algorithms) to process commands.
- For Explicit or Implicit Objectives: Explicitly, it performs tasks like turning on lights or setting alarms. Implicitly, it learns your preferences, such as adjusting lights to your favourite settings at specific times.
- Infers, from the Input it Receives: When you say, "Echo, play my workout playlist," it interprets your voice command, identifying the key intent and matching it to your saved playlist.
- How to Generate Outputs: It responds by playing the desired music, providing weather updates, or even adjusting the room temperature based on additional inputs.
- That Can Influence Physical or Virtual Environments: Adjusting the thermostat or turning off lights affects your physical environment, while curating playlists or syncing calendars impacts your virtual space.
- Levels of Autonomy and Adaptiveness: Over time, it learns your habits, like suggesting workout playlists at your usual exercise times or offering reminders based on recurring tasks.

By breaking down AI's definition and exploring a real-world example, we can see how its building blocks come together to create systems that are both functional and impactful. Moreover, we can better understand the range of AI capabilities and consider the implications for ethics, governance, and society. Understanding these principles lays the foundation for exploring AI's technical and societal dimensions in the sections ahead.

Exploring AI: The Essentials

What made modern AI possible?

The breakthroughs in AI over the past decade, particularly in Deep Learning (DL) and Large Language Models (LLM), have been driven by three key enablers:

1. **Powerful Processing Capabilities:** Advances in hardware, such as Graphics Processing Units (GPUs), have significantly sped up machine learning processes, enabling the training of complex models at unprecedented speeds. However, this comes at a significant environmental cost due to the vast consumption of energy, water, and minerals required.

2. **The Availability of Vast Amounts of Data:** The digital era has created an explosion of data from industries, social media platforms, and everyday online interactions. This data serves as the raw material for AI systems to learn and improve.
3. **Optimised Data Preparation through Microtasking:** Platforms like Amazon Mechanical Turk have allowed organisations to rely on distributed, low-cost labour to classify, label, and annotate data, fueling the development of large datasets needed for AI training.

While these factors have propelled innovations in language processing, image recognition, generative AI, and autonomous systems—they have come at significant ethical, social, political, and environmental costs.

Two broad categories of AI

AI can be broadly classified into two categories that define its current and aspirational capabilities:

1. **Narrow AI (Weak AI):** is the most common form of AI today. Narrow AI is designed to perform specific tasks with remarkable accuracy. For example, Narrow AI can play chess, drive cars, recognise speech or text, predict weather, filter spam, power virtual assistants, or recommend content. These systems, however, excel at their specialised tasks but cannot transfer their learning to other tasks. A self-driving car, for instance, cannot learn to play chess.
2. **General AI (Strong AI/Artificial General Intelligence - AGI):** this is an aspirational goal and vision in the AI field. General AI refers to systems that can understand, learn, and apply knowledge broadly, similar to human intelligence. General AI would be able to perform a variety of tasks, adapt to new situations, and apply its learning across domains. General AI would be able to perform a variety of tasks, adapt to new situations, and apply its learning across domains.

What AI can do and where it falls short?

To better understand what AI excels at and its limitations, Arvind Narayanan, a computer science researcher, proposed a framework categorising AI technology into three groups:

Category	Description	Examples
AI-driven technologies with genuine and rapid progress	AI technologies in this category show rapid and reliable advancements, excelling in pattern recognition and structured tasks.	Content identification (e.g., Shazam, reverse image search), face recognition, medical diagnosis from scans, speech-to-text, deep-fakes, etc..
AI-driven technologies that are far from perfect, but improving	In this category AI struggles with tasks requiring nuanced understanding or ethical judgment, but there is steady progress.	Spam detection, detection of copyrighted materials, hate speech detection, etc..
Fundamentally dubious AI-driven technologies	These technologies attempt to predict complex, context-driven human behaviours or outcomes, often leading to ethical and societal concerns.	Predicting criminal recidivism, job performance, predictive policing, identifying “at-risk” children.

The beating heart of AI: Algorithm, data, model

At the core of every AI system are three building blocks—Algorithms, Data, and Models—that work together like a recipe, ingredients, and final dish to help machines learn and make decisions.

- **Algorithm** is a set of step-by-step instructions that tells a computer how to perform a task. Think of it like a recipe for making a sandwich: you decide how many slices of bread to use, which ingredients to layer first, and when to toast it. In AI, an algorithm is the core process that the computer follows to learn from data and make decisions or predictions.
- **Data** refers to the raw information that we feed into our algorithms. In the sandwich example, data could be surveys of people's favourite ingredients, lists of common allergies, or ratings of which sandwich combinations taste best. The quality and quantity of data directly affect how well an AI system can "learn."
- **Model** is the mathematical representation of a problem that an AI system trains on using data. After we feed data into our algorithm, the algorithm produces (or updates) a model. For our sandwich scenario, the model might be a set of rules or patterns that indicate, for example, "If people like cheese and tomatoes, they often also like lettuce." Once this model is built, it can predict new sandwich combinations that people might enjoy.

The driving force of AI: The main techniques

The driving force behind contemporary AI systems lie several core techniques that empower machines to learn from data, make decisions, and perform tasks that would typically require human intelligence, such as speech and image recognition or language translation. While these techniques are great in number and too specialised to explore in depth here, we will provide a brief overview of some of the main AI techniques.

Machine Learning

Machine Learning (ML) is a technique of AI focused on teaching computers to learn from data and improve over time without being explicitly programmed for every scenario. Machine-learning-based computational approaches made possible many advances in the field of AI, such as: face recognition, autonomous vehicles, natural language processing, and more. For example, by analysing thousands of emails, a machine learning system can learn to distinguish between spam and non-spam messages. Essentially, it's about enabling machines to gain insights and make predictions based on their analysis of data, thereby performing tasks that would typically require human intelligence. There are three main ML approaches:

- **Supervised Learning**: The data is labelled. Imagine you have a labelled list of sandwiches and their popularity ratings. You tell the AI which sandwiches are "good" or "bad" based on how many people liked them. The AI then looks for patterns—like which ingredients appear in the most popular sandwiches—and uses that to predict new, tasty combinations.
- **Unsupervised Learning**: The data is not labelled. Instead, the AI looks for hidden patterns or groupings on its own. If you just provide a bunch of ingredient lists without telling it which

sandwiches are popular, it might group sandwiches by common characteristics (e.g., “spicy sandwiches,” “vegetarian sandwiches,” “cheese-based sandwiches”), potentially helping you discover entirely new categories.

- **Reinforcement Learning**: The AI learns by trial and error and through rewards or penalties. If you have a robotic sandwich maker, every time it successfully assembles a popular sandwich, it receives a reward (like a higher score). If it picks a terrible ingredient combo, it gets a penalty. Over time, it figures out the best sequence of actions to maximise its reward (i.e., making the best sandwiches more often).
  A specialised subset of reinforcement learning is **Reinforcement Learning with Human Feedback (RLHF)**. In this method, human reviewers evaluate and score the model’s initial outputs based on criteria such as accuracy, precision, relevance, or bias. The model is then fine-tuned using this feedback, giving preference to higher-scoring outcomes. Positive rewards are assigned to good outcomes, while penalties are applied to poor ones, enabling the model to improve over time. For instance, RLHF can enhance a fraud detection system, helping it become more precise in identifying suspicious transactions.

Artificial Neural Networks

An Artificial Neural Network (ANN) is an AI system inspired by the human brain’s network of neurons. It takes in data, processes it through many interconnected “neurons” or units, and learns to recognize complex patterns. If you fed sandwich images or ingredient lists to a neural network, it would learn which features make a sandwich appealing. A prominent example of ANNs is Google's AlphaGo. In 2016, this advanced AI system made headlines when it triumphed over the world's top Go player, demonstrating the potential of ANNs in mastering complex tasks. Other examples which employ ANNs are facial recognition technology used in smartphones or security systems, identifying individuals' faces with great accuracy.

Deep Learning

Deep Learning (DL) is a specialised subset of ANNs that mimics the workings of the human brain in processing data and creating patterns for use in decision making. It's called "deep" because it makes use of deep ANNs—layers upon layers of interconnected nodes, or artificial neurons, much like the neural networks in our brains. These layers can learn to recognise complex features in data, from the simplest elements in the early layers to highly complex patterns in the deeper layers. In other words, all DL models are ANNs, but not all ANNs qualify as ‘Deep’—the key difference is the number of layers and how they process data. To use the analogy of sandwich—imagine stacking several taste-testing teams, each focusing on increasingly detailed sandwich features. One layer might identify the bread, another layer the fillings, and yet another how the ingredients are arranged. By stacking multiple layers (deep networks), the system can uncover more complex patterns—like not just which bread pairs well with turkey, but also whether the colour, texture, or order of ingredients affects people’s preferences.

Within Deep Learning, there are several specialised architectures, each suited to different types of problems:

- **Deep Neural Networks (DNN):** These have multiple hidden layers, each refining the patterns identified by the previous layer. Think of it like a series of "checks" for your sandwich: first, identifying the bread, then the main filling, then toppings, and so on.
- **Recurrent Neural Networks (RNN):** These are designed for sequential data like text or time-series. In sandwich-making, an RNN might help the AI consider the order of ingredients (e.g., bread first, then lettuce, then cheese…).
- **Convolutional Neural Networks (CNN):** These are specialised for image data. If you have pictures of sandwiches with various ingredients, a CNN can learn which visual features (e.g., color, shape) correspond to the best-tasting options.
- **Generative Adversarial Networks (GAN):** A GAN pairs two networks—a Generator (which creates new data) and a Discriminator (which tries to detect what's real vs. fake). The Generator might create new sandwich recipes, while the Discriminator tries to judge how "convincing" or tasty these recipes are. Over time, the Generator gets better at making truly innovative (and hopefully delicious or maybe unforeseen before) sandwich ideas. (Note: the same technique behind these creative sandwich combinations also powers "deepfakes," realistic but potentially misleading or harmful images and videos—so it's important to be aware of the ethical implications.)

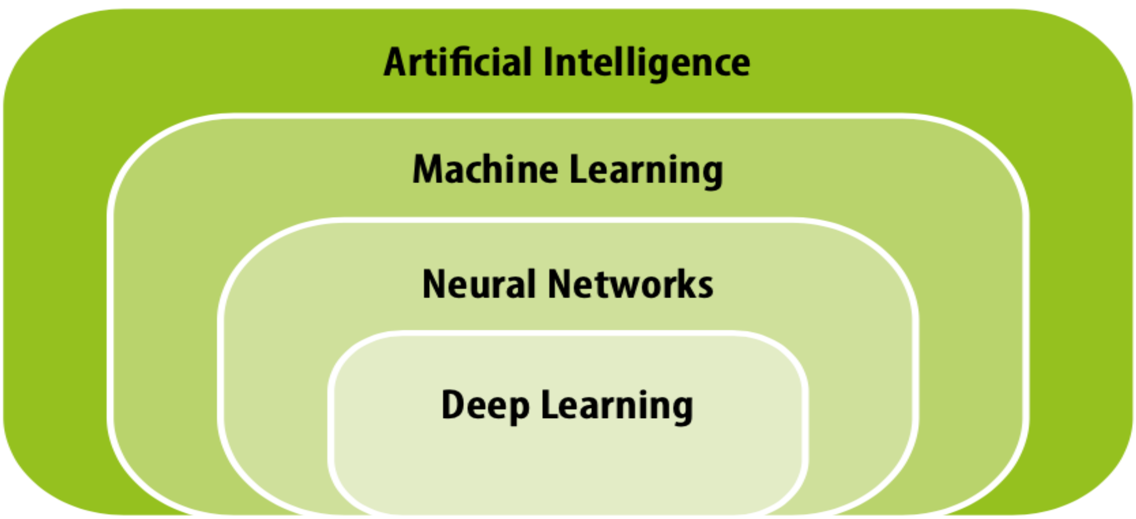

Figure 2.1 The relationship between Artificial Intelligence, Machine Learning, Neural Networks and Deep Learning © UNESCO "AI and education: Guidance for policy-makers"

From techniques to reality: AI in our world

Now that we've explored the beating heart of AI—its core building blocks (Algorithm, Data, Model) and central techniques (Machine Learning, Deep Learning, and specialised architectures)—let's see how these come to life in tools and services we encounter daily. Below we present an overview of the key AI-driven technologies.

Generative AI and Large Language Models (LLMs)

Generative AI is a specialised subfield of AI that focuses on creating new content—whether it's text, images, music, code, or even virtual environments. Within this subfield, LLMs are advanced neural networks trained on enormous amounts of text data. They learn patterns in language so well that they can produce human-like responses, write essays, translate between languages, and more. They're built on DL techniques, which fall under the broader umbrella of ML. These models power many cutting-edge applications in chatbots, art generation, and more.

Examples

- AI-driven chatbots like ChatGPT, Bard, Gemini, etc. that respond naturally, assist with writing, or answer questions.
- Tools like DALL·E or Midjourney produce imaginative artwork from text prompts.
- Platforms like AIVA or Amper Music compose original music, from classical scores to modern pop tunes, often indistinguishable from tracks written by human composers.
- WaveNet is a generative model for raw audio, primarily used for speech synthesis, can produce highly realistic voices and soundscapes, opening avenues in voice cloning and audio design.

Natural Language Processing (NLP)

NLP teaches computers to understand, interpret, and generate human language, whether in written or spoken form. By recognising grammar, context, and semantics, NLP-driven systems can perform tasks like translation, summarisation, or sentiment analysis. NLP spans several AI techniques—some simpler systems may rely on classical machine learning, while more advanced ones often use Deep Learning or LLMs (like ChatGPT) to handle complex language tasks.

Examples

- Voice Assistants like Siri, Alexa, Google Assistant, etc., convert speech to text and respond to spoken commands.
- Language Translation apps like Google Translate or Deepl instantly convert text between different languages.
- Sentiment Analysis tools monitor social media or customer reviews to gauge public opinion in real time.

Computer Vision and Image Recognition

Computer Vision enables machines to see and interpret images or videos, much like the human visual system. AI can detect objects, recognise faces, and classify images by analysing pixel patterns.

Computer Vision often relies on DL methods—especially CNNs—to process visual information effectively.

Examples

- Facial Recognition Technologies like Apple Face ID which unlocks iPhones by scanning the user's face.
- Medical imaging technologies such as Aidoc or IBM Watson Health assists radiologists by analysing medical scans to detect anomalies or potential health risks; or examine CT scans to help identify urgent medical conditions such as strokes or pulmonary embolisms.
- Wildlife monitoring often uses camera traps, which automatically identify animal species in remote locations. For instance, 'Conservation AI' harnesses machine learning for various conservation projects.
- Urban cleanliness and environmental monitoring is exemplified by Amsterdam's Smart Litter Detection system, which uses AI-driven camera systems to spot garbage or litter on sidewalks and streets. The system sends alerts to cleanup teams, reducing manual inspections and improving waste management efficiency.

Multimodal AI

Multimodal AI refers to systems that can process and combine different types of data—like text, images, audio, and video—to perform complex tasks. By linking modalities, these models better understand context and solve problems that rely on multiple sources of information. For example, they can analyse an image, read associated text, and generate a meaningful description. We can locate multimodal AI at the intersection of NLP, computer vision, and sound analysis. It uses advances in DL (e.g., transformers, CNNs) to connect different data types into a cohesive understanding.

Examples

- GPT-4 can process text and images, allowing users to ask questions about photos, diagrams, or charts in addition to text.
- Accessibility tools such as Seeing AI by Microsoft assist visually impaired individuals by describing their surroundings, recognising text in photos, or identifying objects in real-time.
- Numerous cities around the world utilise AI-enhanced microphones and cameras to monitor noise pollution levels, such as loud engines or construction noises, while cameras provide the corresponding visual evidence for enforcement.

Speech Recognition and Sound Analysis

This area involves AI systems that process and interpret audio data, whether speech or general sounds (like music, gunshots, or environmental noises). Speech recognition typically combines NLP (to understand the words being said) with DL (to convert sounds into text). Sound analysis also uses pattern recognition methods to classify or detect specific audio signals.

Examples

- Music identification apps like Shazam match short audio clips to vast song databases.
- AI-driven tools like Rainforest Connection's Guardian detect chainsaw noises in forests to prevent illegal logging or gunshots in urban areas to alert law enforcement for rapid response.
- Speech-to-text features, such as Otter.ai or Microsoft Azure Speech Services, assist individuals with hearing impairments in daily communication by transcribing spoken words into written text in real time.

Predictive Analytics and Recommendation Systems

Predictive Analytics uses AI to forecast future trends or outcomes based on historical data. Recommendation Systems tailor suggestions for users (e.g., which movie to watch, which product to buy) by analysing their past behaviour and preferences. These systems commonly use ML approaches (both classical and deep-learning-based) to analyse patterns in large datasets. They're integral to many consumer-facing platforms.

Examples

- Streaming services such as Netflix, YouTube, or Spotify, personalise recommendations based on viewing/listening history.
- E-commerce firms suggest products by comparing your shopping habits with those of similar users.
- Weather forecasting tools process huge amounts of atmospheric data to predict short- and long-term weather patterns.

Autonomous Systems

Autonomous systems combine multiple AI approaches to operate independently with minimal human intervention. They make decisions on the fly, adapting to new situations. They often use a fusion of ML (to learn from data), Computer Vision (to understand surroundings), and sensor inputs (like radar or Light Detection and Ranging) to interact safely with dynamic environments.

Examples

- Self-driving cars integrate sensor data, vision, and ML to navigate roads and traffic.
- Robotic process automation handles repetitive tasks—like sorting packages or inventory management.
- Drones and delivery robots use onboard AI to deliver goods, capture aerial imagery, or perform inspections with minimal supervision.

We have seen how vast and complex the world of AI truly is—spanning far beyond tools like ChatGPT, which often dominate public imagination. Despite its powerful capabilities, AI is not infallible; it remains fragile, brittle, and capable of causing real harm in society. In the next section, we will explore this critical and human dimension of AI, examining its impact and implications on human rights, democracy, education, and other areas.