



Starter Kit for Ethical AI

Ethical AI Fundamentals Guide

www.valuesai.eu



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Ethical AI Fundamentals Guide © 2025 by VALUES is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by-sa/4.0/)

Introducing the VALUES Project

The **VALUES project** is a forward-thinking initiative aimed at empowering young people across Europe by providing them with the essential skills and ethical understanding needed to navigate the rapidly developing world of artificial intelligence (AI). As AI technologies become an integral part of everyday life, it is essential for youth to not only acquire technical knowledge but also to develop a strong awareness of the ethical considerations surrounding AI. The project's core objective is to equip young people, particularly those facing employment challenges, with the knowledge and tools to critically evaluate AI technologies, ensuring they can engage with these systems responsibly. By fostering a deeper understanding of AI ethics, the VALUES project aims to prepare young people to become responsible digital citizens and ethical innovators, ready to contribute positively to the digital economy. The anticipated outcomes include increased AI literacy, a more ethical approach to technology use, and enhanced employment opportunities in the growing AI sector.

What is the VALUES Starter Kit for Ethical AI?

*The **VALUES Starter Kit for Ethical AI** is your gateway to exploring the fascinating and essential world of ethical AI. Designed for youth workers and educators, this dynamic resource offers you the tools to dive into the complexities of AI, from its core principles to the ethical dilemmas it presents. Packed with real-world examples, practical strategies, and*

engaging activities, this kit will help you spark meaningful discussions with young people about the impact of AI on their lives. Whether you're introducing AI for the first time or guiding young people through its ethical implications, this Starter Kit will empower you to teach with confidence and ignite curiosity about the responsible use of AI in today's world.

contents

01	INTRODUCTION TO AI AND ETHICS
02	CORE PRINCIPLES OF ETHICAL AI
03	ETHICAL DILEMMAS IN AI
04	GUIDING YOUTH THROUGH ETHICAL AI DISCUSSIONS
05	ANNEXES



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Ethical AI Fundamentals Guide © 2025 by VALUES is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by-sa/4.0/)

01

INTRODUCTION TO AI AND ETHICS



Introduction to AI and Ethics

What is AI?

Artificial Intelligence (AI) refers to the field of computer science focused on creating systems that can perform tasks traditionally requiring human intelligence. These systems are designed to simulate human cognitive functions such as learning, reasoning, problem-solving, perception, and language understanding. The aim of AI is not just to replicate human tasks but to enable machines to improve themselves over time through experience and data. In today's world, AI is rapidly transforming various industries, and understanding its key concepts is essential for youth workers, educators, and young people to navigate this evolving landscape.

Let's explore some of the core concepts within AI:

1. Machine Learning (ML)

Machine Learning is one of the most prominent subfields of AI, where algorithms enable computers to learn from and make predictions based on data without being explicitly programmed. Essentially, in machine learning, the system improves its performance by processing large amounts of data and identifying patterns within that data.

For example, machine learning is at the core of recommendation systems used by platforms like Netflix and YouTube. These systems learn from users' previous actions (such as viewing history) to suggest movies or videos that they are likely to enjoy.

Real-World Example:

- **Netflix Recommendations:** Netflix uses machine learning to suggest movies or series based on your viewing history, preferences, and ratings. The more data the system collects about your behaviour, the better it becomes at predicting what you'll want to watch next.

2. Natural Language Processing (NLP)

Natural Language Processing enables computers to understand, interpret, and respond to human language in a way that is both meaningful and useful. NLP helps machines process language in various forms, such as spoken language (speech recognition) or written text (text analysis). NLP can be used in applications such as chatbots, virtual assistants, and translation tools.

One key challenge in NLP is teaching machines to understand context, tone, and meaning in human language. This allows computers to not only interpret words but also grasp the underlying intent of a sentence or phrase.

Real-World Example:

- **Voice Assistants (Siri, Alexa):** These devices use NLP to process spoken commands and respond appropriately. For example, when you ask Alexa to play a song, NLP helps it understand your request and execute the task accurately.



WATCH: “Natural Language Processing In 5 Minutes”

Introduction to AI and Ethics

What is AI?

3. Deep Learning

Deep Learning is a subset of machine learning that uses artificial neural networks to simulate the way the human brain processes information. These networks consist of multiple layers of nodes (hence the term “deep”), allowing the system to learn and make decisions with a high level of accuracy.

Deep learning is particularly effective in tasks like image recognition, where the system learns to identify objects, faces, or scenes in pictures. The more data the system processes, the better it becomes at recognising patterns and making predictions.

Real-World Example:

- **Facial Recognition Technology:** Deep learning algorithms power facial recognition systems used in security cameras, smartphones, and social media platforms. These systems can identify individuals in photos or videos, enhancing security and personalisation.

4. Computer Vision

Computer Vision is an area of AI that enables machines to interpret and understand visual information from the world around them. This includes tasks like identifying objects in images, recognizing faces, or even interpreting videos. Computer vision has applications in a wide range of industries, including healthcare, retail, and transportation.

For instance, self-driving cars rely heavily on computer vision to detect obstacles, traffic signals,

and pedestrians on the road. By processing real-time images from cameras installed on the vehicle, the car can make decisions about its movements.

Real-World Example:

- **Self-Driving Cars:** Companies like Tesla and Waymo are using computer vision to enable autonomous driving. The cars use AI algorithms to process data from cameras and sensors, helping them navigate streets and make decisions such as stopping at traffic lights or avoiding obstacles.

5. AI in Everyday Life

While the above concepts seem technical, AI is already woven into the fabric of our everyday lives. From social media to online shopping, AI technologies help make decisions, improve experiences, and automate tasks. Here are a few key examples of how AI is used:

- **Social Media Algorithms:** AI powers the algorithms that determine what posts, ads, and content you see on platforms like Facebook, Instagram, and Twitter. These systems analyse your behaviour and preferences to curate a personalised feed.
- **Online Shopping:** Retailers like Amazon use AI to recommend products based on your previous purchases or search history. This recommendation engine uses machine learning to predict what you're likely to buy next.
- **Healthcare:** AI is also making strides in the medical field, assisting doctors in diagnosing diseases like cancer, predicting patient outcomes, and providing personalised treatment plans based on data analysis.

Introduction to AI and Ethics

The Role of AI in Society

Artificial Intelligence is increasingly becoming an integral part of our daily lives, shaping how we interact with the world, how we work, and how we engage with each other in society. AI is not just confined to technology labs or large corporations; it has found its way into social platforms, the workplace, and even civic engagements, influencing and transforming various sectors and areas of life. This section delves into AI's role in society, with a focus on its impact on social life, careers, and youth engagement.

1. Common Uses in Social Life

AI is now ubiquitous in social life, and its influence is deeply felt in many areas, particularly through social media and entertainment technologies. Here are some of the most common ways AI shapes social experiences:

Social Media Algorithms

Social media platforms like Facebook, Instagram, TikTok, and Twitter use AI-powered algorithms to curate and personalise content for users. These algorithms assess your interactions—likes, shares, comments, and the time spent on posts—to determine what content is most likely to capture your attention next. The more you engage with specific types of content, the more the platform's AI tailors your feed to suit your preferences.

These algorithms are designed to increase user engagement, which leads to longer screen time and, consequently, more opportunities for advertisers to target users with personalised ads. However, this process also raises questions about privacy, data collection, and the potential for users to be manipulated by targeted content, including misinformation and sensationalism.

Real-World Example:

- **Instagram's Explore Page:** The AI behind Instagram's Explore page predicts which posts you might find interesting based on your past likes and engagement patterns. If you often like photos related to fashion, Instagram's algorithm will show you more fashion-related content, customising your experience to keep you engaged.

Deepfakes

Deepfakes are AI-generated synthetic media that superimpose images or videos onto existing footage, often creating hyper-realistic fake videos. These videos are commonly used to impersonate someone, making it look like they said or did something they never actually did. Deepfake technology uses AI techniques, particularly deep learning, to create these convincing manipulations, which can range from harmless entertainment to harmful content intended to deceive or manipulate.

The rise of deepfakes poses serious ethical concerns, especially when they are used to spread misinformation, defamation, or political manipulation. They challenge the credibility of media sources and can undermine trust in visual content.

Real-World Example:

- **Deepfake Videos in Politics:** Deepfakes have been used in political campaigns to create fabricated speeches or events involving public figures, potentially influencing elections or public opinion. The ability to manipulate video content raises crucial questions about how we authenticate information and trust media sources.

Introduction to AI and Ethics

The Role of AI in Society

2. The Impact of AI on Careers

AI's influence on the workforce is profound, and it is reshaping many industries, creating new job opportunities while simultaneously rendering certain jobs obsolete. As AI becomes more sophisticated, the demand for AI-related skills has grown exponentially, making it a critical area of focus for young people entering the workforce.

Automation and Job Displacement

One of the most discussed effects of AI is automation—the use of machines and algorithms to perform tasks traditionally carried out by humans. While automation increases efficiency, it also threatens job security in various sectors, particularly in manufacturing, transportation, and data entry. AI-driven robots and algorithms are already replacing repetitive tasks that once required human labour, resulting in job displacement for workers in those fields.

Real-World Example:

- **Self-Checkout Machines:** Many retail stores now use self-checkout machines, which reduce the need for cashiers. As these machines improve, they may replace many jobs in the retail sector, making it crucial for workers to upskill and learn to work alongside automated systems.

The Growing Demand for AI-Related Skills

On the flip side, AI is creating new job opportunities in fields like data science, machine learning, robotics, and AI ethics. As industries continue to integrate AI technologies, there is an increasing demand for individuals with expertise in these areas. For young people, acquiring skills in AI and related fields can significantly enhance career prospects, ensuring they are well-equipped to thrive in the future job market. AI is not limited to the tech industry. Fields such as

healthcare, law, finance, and marketing are also incorporating AI technologies. For example, in healthcare, AI is being used to predict patient outcomes, assist in diagnostic imaging, and personalise treatment plans, creating new roles for AI specialists in these areas.

Real-World Example:

- **AI in Healthcare:** Companies like IBM Watson Health use AI to analyse medical data and provide insights that help doctors make more accurate diagnoses. These advancements require skilled professionals to manage and implement AI systems in healthcare settings.

READ: The Impact of AI on Job Roles, Workforce, and Employment



Introduction to AI and Ethics

The Role of AI in Society

3. AI's Relevance to Youth (Social Life, Career, Civic Engagement)

AI's role extends far beyond the workplace and social media—it also has profound implications for how young people engage with their communities and the world. The growing integration of AI into various aspects of life means that youth need to be informed and equipped to engage with these technologies in meaningful ways, both socially and civically.

Social Life and Civic Engagement

AI is influencing how young people engage with the world around them. From social media algorithms to AI-driven tools used in activism, AI plays a central role in shaping how information is disseminated and consumed. Understanding AI helps young people critically engage with the content they encounter online, make informed decisions about their digital footprint, and avoid falling victim to manipulation, whether from advertising or misinformation.

AI is also increasingly used in civic activism, where it helps amplify the voices of youth in important societal issues, such as climate change, human rights, and education. For example, AI-powered platforms can help activists track and share real-time information about protests, mobilise supporters, and analyse social trends that can inform activism strategies.

Real-World Example:

- **AI in Social Movements:** Platforms like

Change.org use AI to help organise petitions and mobilise people around social causes. AI can analyse social media trends to identify key issues and bring awareness to important topics.

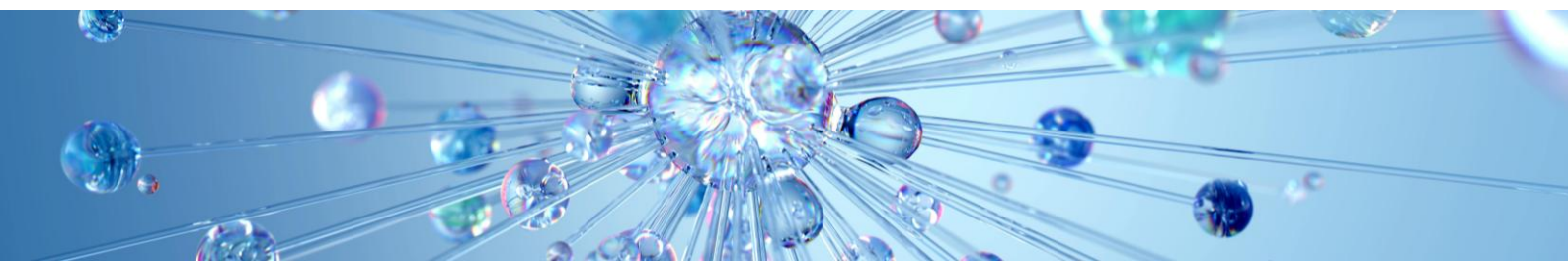
Career and Education

For young people entering the workforce, AI is becoming an essential part of many fields. As previously mentioned, industries ranging from healthcare to finance to entertainment are increasingly reliant on AI technologies. By understanding AI, youth can better prepare themselves for careers in these sectors. Moreover, AI-related skills are becoming increasingly important in education. Schools and universities are integrating AI into their curriculums to teach students not only about AI itself but how to think critically about its impact on society.

Youth workers and educators have a key role to play in preparing the next generation to use AI responsibly and innovatively. Equipping young people with AI literacy is no longer just about giving them technical skills; it's about empowering them to be active, informed, and ethical participants in a society where AI plays a growing role.

Real-World Example:

- **AI and Personalised Learning:** In education, AI is being used to create personalised learning experiences. Platforms like Khan Academy use AI to tailor learning materials based on individual student progress, ensuring each learner receives content suited to their needs.



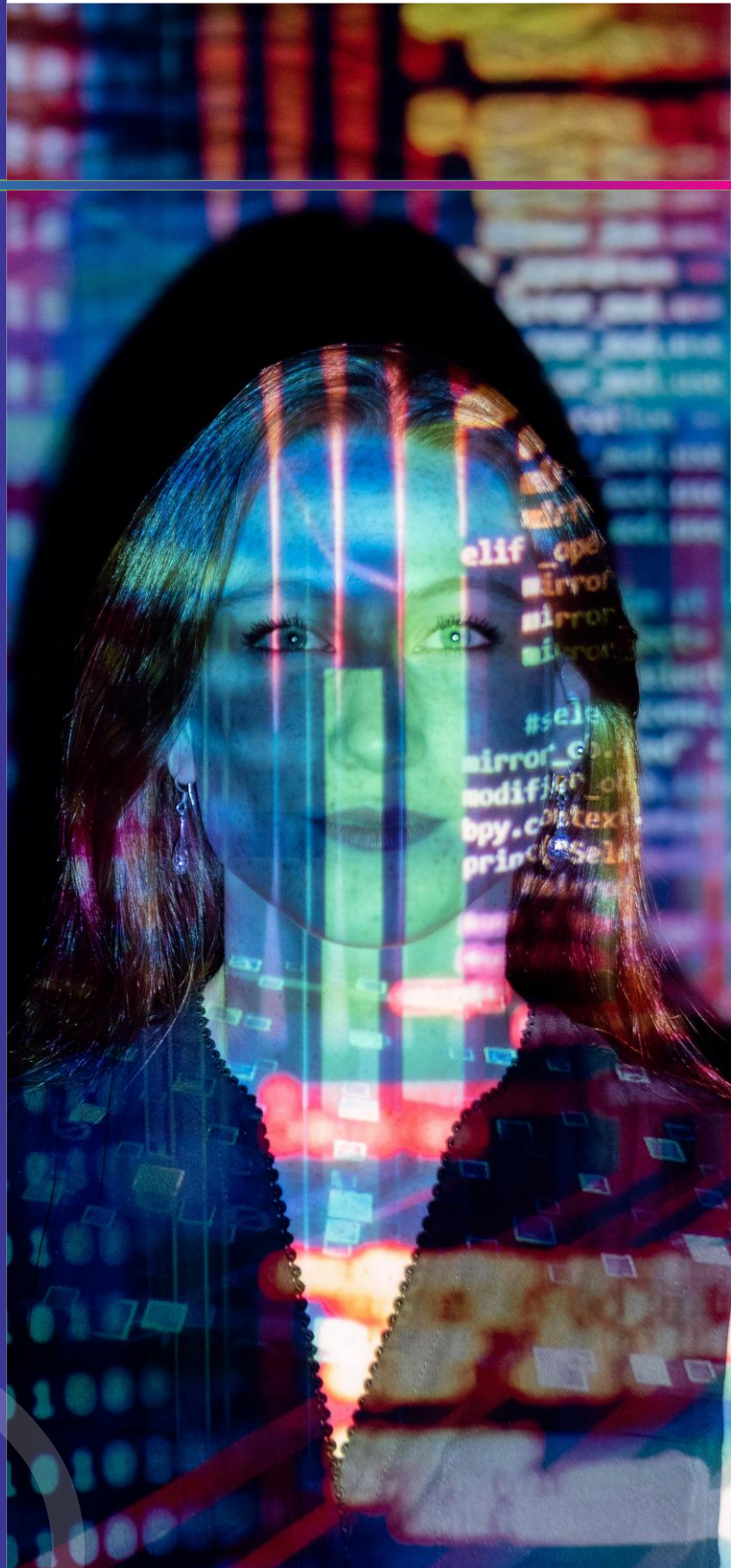
Conclusion:

Introduction to AI and Ethics

As we've seen in this chapter, Artificial Intelligence is more than just a technological innovation—it's a powerful force shaping our social lives, careers, and civic engagement. AI is already embedded in the fabric of our everyday experiences, from the social media algorithms that curate our feeds to the ways it influences the job market, creating new opportunities while also presenting new challenges. For young people, AI is not just a concept for the future—it's a present reality that touches many aspects of their lives and will play an even greater role as they step into the workforce and engage with the world around them.

The ethical implications of AI are profound and far-reaching. Whether it's ensuring transparency in algorithms, addressing bias, or protecting personal data, the responsibility to use AI ethically falls on all of us. This chapter has introduced you to the foundational concepts of AI and its core ethical principles, setting the stage for the deeper discussions ahead. Understanding these concepts is the first step in helping young people navigate the complex AI landscape responsibly, making informed decisions that align with ethical guidelines and European values.

As youth workers, educators, and trainers, it is essential that we equip young people with both the technical skills and the critical thinking abilities needed to engage with AI technologies. The next chapters will delve deeper into the specific ethical dilemmas AI presents and offer practical tools to guide youth through discussions on these issues, ultimately preparing them to be not only skilled users of AI but also responsible and ethical participants in an AI-driven world.





“AI is not neutral. It's not just a tool; it's a reflection of the values of the people who create it.”

Timnit Gebru,
AI researcher and advocate
for ethical AI



02

CORE PRINCIPLES OF ETHICAL AI



Core Principles of Ethical AI

What is Ethics and Why Does it Matter?

Ethics is the branch of philosophy that deals with questions about what is morally right and wrong, good and bad, fair and unfair. In the context of artificial intelligence (AI), ethics is crucial because AI systems can influence critical aspects of human life, from social interactions to employment opportunities, healthcare, and even personal freedom. As AI systems increasingly shape our world, it is essential to ensure that they are designed, developed, and used in ways that align with fundamental moral principles, including fairness, justice, and respect for individuals' rights.

Ethics in AI helps to navigate the complexities of how these technologies should interact with people and society. AI technologies often make decisions or suggestions based on vast amounts of data, which can lead to unintended consequences if the systems are not carefully designed with ethical guidelines in mind. For example, AI algorithms may inadvertently reinforce stereotypes, discriminate against certain groups, or invade individuals' privacy. To prevent these negative impacts, ethical principles must be integrated into every phase of AI development, from design and data collection to implementation and use.

The importance of ethics in AI is especially relevant as AI systems become more autonomous and influential. Without ethical oversight, AI systems could perpetuate inequality, exacerbate existing biases, or make decisions that violate fundamental human rights. For example, AI-based systems used in hiring processes may unintentionally favour certain groups over others if not properly monitored for bias. Similarly, AI technologies used in surveillance could infringe on privacy if not regulated appropriately.

Ethics in AI is also necessary to build trust between people and these technologies. If users are unsure whether AI systems are being used responsibly, they may become resistant to adopting these technologies, even if they have the potential to offer significant benefits. In contrast, if AI systems are seen as ethical, fair, and transparent, individuals and society are more likely to embrace their use. By embedding ethical principles into AI, we can create systems that are not only effective and efficient but also equitable and just.

Ultimately, ethical considerations in AI are not just about avoiding harm—they are about ensuring that AI systems are designed and used in ways that positively contribute to society. AI ethics involves making thoughtful decisions about how technology can be used for the common good while safeguarding individual rights and freedoms. Ethical AI encourages us to think critically about how we want these systems to impact our world and how to build technology that aligns with the values we hold dear.

As we move forward into an increasingly AI-driven future, understanding and addressing ethical issues is key to ensuring that these technologies serve humanity in a way that is fair, just, and aligned with our shared moral compass.



Core Principles of Ethical AI

Bias in AI Algorithms

AI systems make decisions based on data, but when the data is biased, the results can be biased too. Bias in AI refers to unfair or unequal outcomes that disadvantage certain groups. This can occur at different stages, particularly in data collection, training, and application.

How Bias Arises in AI

Bias emerges when AI systems are trained on data that is incomplete, unrepresentative, or influenced by historical biases. For example, facial recognition systems often perform poorly on women and people of colour because the training data used is predominantly white. Similarly, AI used in hiring may perpetuate biases in past recruitment decisions, unfairly disadvantaging women or minority groups.

Types of Bias in AI

- **Selection Bias:** Occurs when data used is not representative of the population.
- **Label Bias:** Happens when the labels in training data are inaccurate or biased.
- **Measurement Bias:** Arises from inaccurate or biased data collection.
- **Aggregation Bias:** Occurs when AI generalizes characteristics of diverse groups as homogeneous.

Addressing Bias in AI

1. **Diverse and Representative Data:** Ensuring that training data represents all demographic groups.
2. **Bias Audits:** Regularly testing AI systems to identify and correct biases.
3. **Transparent Algorithms:** Creating explainable algorithms to understand and rectify biased decisions.

4. **Algorithmic Fairness:** Embedding fairness criteria during the design phase.
5. **Human Oversight:** Involving humans in decision-making, especially in critical applications.

Ethical Implications

Bias in AI can perpetuate social inequalities, infringe on people's rights, and undermine trust in technology. For example, biased hiring algorithms may exclude qualified candidates based on gender or race, while biased criminal justice algorithms can result in unfair sentencing. Transparency and accountability are key to preventing these ethical issues and building trust in AI systems.

Conclusion

To create ethical AI systems, it is crucial to identify, understand, and mitigate bias. By ensuring fairness and justice, AI technologies can better serve all people, regardless of their background or identity.



WATCH: "Can we stop AI from inheriting our biases?"

Core Principles of Ethical AI

Transparency in AI Systems

Transparency in AI refers to understanding how AI systems make decisions. As AI integrates into fields like healthcare and criminal justice, transparency is crucial for trust, accountability, and fairness.

Why Transparency Matters

Transparency ensures that stakeholders, such as developers, users, and regulators, can understand how AI systems operate. It allows users to challenge decisions and fosters trust. This is especially important in high-stakes areas like hiring, lending, and law enforcement.

Key Elements of Transparency

1. **Explainability:** The ability for humans to understand AI decisions, crucial in areas like healthcare and law.
2. **Accessibility of Information:** Making algorithms, data sources, and models clear and open to scrutiny.
3. **Auditability:** Allowing external audits to check AI systems for fairness, bias, and risks.

Challenges to Transparency

- **Algorithm Complexity:** Many AI models, like deep learning, are hard to fully explain due to their "black box" nature.
- **Proprietary Models:** Some companies keep models secret for commercial reasons, limiting transparency.
- **Technical Constraints:** Some systems may not be fully transparent due to technical limitations.

Benefits of Transparency

1. **Trust:** Transparent AI builds confidence by showing how decisions are made.
2. **Accountability:** Transparency helps hold individuals accountable when AI decisions go wrong.
3. **Fairness:** It allows for identification and correction of biases, ensuring fairness.

Strategies for Enhancing Transparency

1. **Clear Documentation:** Providing accessible explanations of how AI models work, including data and training processes.
2. **Open Algorithms:** Sharing source code or model details for external inspection.
3. **AI Impact Assessments:** Assessing AI's social, economic, and cultural impact and making results public.
4. **Stakeholder Engagement:** Involving affected individuals in the AI development process, particularly in sensitive areas.

Conclusion

Transparency is key to ethical AI. It enables understanding, builds trust, and ensures accountability. As AI continues to evolve, prioritising transparency will help guarantee that AI technologies are fair, understandable, and trustworthy.



Core Principles of Ethical AI

Accountability in AI

Accountability in AI refers to who is held responsible when AI systems make mistakes or cause harm. As AI becomes more integrated into everyday life, establishing clear accountability is crucial to ensure ethical usage and prevent abuse.

Why Accountability Matters

AI systems can make decisions that impact people's lives, from job applications to healthcare treatments. When an AI system fails, whether due to design flaws or misuse, accountability ensures that those responsible for developing, deploying, and overseeing the technology are held answerable for its consequences.

Key Issues in AI Accountability

1. **Who is Responsible?:** If an AI system causes harm, is the responsibility with the developers, the users, or the company that owns the system? This question is crucial in fields like autonomous driving or predictive policing.
2. **Assigning Blame:** In complex AI systems, tracing errors back to a specific cause can be difficult. Determining whether an error is due to flawed design, biased data, or misuse can be complex, making accountability harder to establish.
3. **Shared Accountability:** Since AI development often involves multiple parties (developers, data providers, companies), establishing shared responsibility is key to ensuring that accountability isn't shifted onto one party alone.

Challenges to Accountability

- **Black Box Nature:** Many AI models, especially deep learning, are opaque. If decisions are not explainable, it becomes difficult to identify who or what caused the error.

- **Automated Decisions:** In cases where AI systems autonomously make decisions without human intervention, it becomes challenging to establish accountability. Should the developer be responsible for the AI's actions, or should the AI system itself bear some responsibility?
- **Regulatory Gaps:** Many countries lack laws or regulations specific to AI, which can make it harder to hold individuals or companies accountable when things go wrong.

Solutions to Ensure Accountability

1. **Clear Governance:** Establishing a clear framework for who holds responsibility at each stage of AI development and deployment can help ensure accountability.
2. **AI Audits:** Regular external audits can ensure that AI systems are functioning ethically and within the intended guidelines, helping to identify potential errors or biases.
3. **Transparent Reporting:** Developers and companies should disclose the functioning and decision-making processes of their AI systems, ensuring that they can be held accountable for their outcomes.
4. **Legal and Ethical Standards:** Developing clear legal frameworks around AI usage, especially in high-risk sectors like healthcare, finance, and criminal justice, can ensure that individuals or companies are held accountable when things go wrong.

AI accountability is essential to its ethical use. By assigning responsibility, implementing oversight mechanisms, and ensuring transparency, we can build systems that are both effective and fair. As AI continues to evolve, ensuring accountability will be crucial to maintaining public trust and safeguarding against harm.

- **Data Anonymisation:** Only anonymised or properly pseudonymised data
- **Data Usage and Ownership:** Transparency about who owns, controls, and
- **Data Usage and Ownership:** Transparency about who owns, controls, and
- **Data Retention and Deletion:** Personal data must not be retained indefinitely
- **Data Retention and Deletion:** Personal data must not be retained indefinitely

Core Principles of Ethical AI

Protecting Personal Information

As AI systems become more embedded in daily life, the responsible handling of personal data is a pressing concern. From health monitoring apps to facial recognition in public spaces, AI increasingly relies on sensitive data. This makes it essential to uphold privacy standards and comply with legal frameworks like the EU AI Act and the General Data Protection Regulation (GDPR/RODO). Protecting this data is key to maintaining user trust, preventing misuse, and aligning with European values on data ethics.

Why Privacy Protection Matters

AI systems often need access to personal data to function, including location data, biometric identifiers, browsing history, and social interactions. Improper use of this data can lead to discrimination, identity theft, or wrongful profiling. Under the AI Act, high-risk AI systems must meet strict requirements for data quality, governance, and transparency. Violations of GDPR—such as processing data without a lawful basis, or failing to ensure data subject rights—can result in severe penalties. AI developers and users must therefore ensure that personal data is processed lawfully, fairly, and with a clear purpose.

Key Issues in Personal Information Protection

1. **Data Collection and Consent:** Users must be clearly informed about what personal data is being collected and why. Consent must be explicit and freely given—implied or bundled consent is not valid under GDPR.
2. **Data Security:** AI systems must incorporate strong technical safeguards such as encryption and access controls. Any breach could result in GDPR reporting obligations within 72 hours.
3. **Data Anonymization:** Only anonymised or properly pseudonymised data should be used for training non-personal AI models. However, developers must avoid “re-identifiable” data sets that risk revealing individual identities through inference.

4. **Data Usage and Ownership:** Transparency about who owns, controls, and benefits from the data is vital. Under GDPR, individuals have the right to know how their data is used—even if AI is involved.
5. **Data Retention and Deletion:** Personal data must not be retained indefinitely. AI systems should include mechanisms to delete data once it is no longer necessary, and allow users to exercise their right to erasure (right to be forgotten).

Sensitive data such as racial or ethnic origin, political opinions, religious beliefs, or data concerning children should never be input into AI systems unless there is a lawful basis and appropriate safeguards in place under GDPR. In most educational or youth settings, such data should be strictly avoided.

Ethical Guidelines for Protecting Personal Information

1. **Transparency:** Users should always know what data is collected, how it is used, and whether any profiling or automated decision-making is involved.
2. **Data Minimization:** AI systems should only collect the minimum necessary data for their function—this reduces risk and increases compliance.
3. **User Control:** Individuals must be able to access, correct, or delete their data. This is a core right under GDPR and essential for building ethical AI literacy.
4. **AI Explainability:** Explainable AI should be the default, particularly when decisions affect rights or opportunities. Users have the right to meaningful information about logic involved in automated decisions.
5. **Data Protection by Design:** Developers should build privacy protections into every stage of AI system design—from planning to deployment—ensuring compliance with both GDPR and emerging AI regulations.

Core Principles of Ethical AI

The Ethics of Deep Fakes

Deep fakes are AI-generated media, such as images, videos, or audio, that manipulate reality with remarkable realism. These synthetic creations raise significant ethical concerns due to their potential for misuse. While deep fakes can be used creatively, their ability to deceive, spread misinformation, and infringe on privacy makes them a critical issue in AI ethics.

What are Deep Fakes?

A deep fake is created using deep learning algorithms to superimpose one person's face on another's or alter speech patterns, making it seem like someone is saying something they didn't. The realistic nature of deep fakes has made them a powerful tool for both creative and harmful purposes.

Ethical Concerns Surrounding Deep Fakes

- 1. Misinformation and Manipulation:** Deep fakes can be used to manipulate public opinion or spread false information. For instance, they've been used to impersonate political figures, which could influence elections or alter public perception.
- 2. Consent and Privacy:** Creating deep fakes without someone's consent violates their privacy and autonomy. Individuals' likenesses or voices can be used maliciously, resulting in defamation, harassment, or emotional harm.
- 3. Accountability and Responsibility:** The anonymous nature of deep fake creation complicates accountability. Without clear responsibility, it's difficult to address the negative impacts caused by deep fakes, such as reputational damage or emotional distress.
- 4. Erosion of Trust:** As deep fakes become more sophisticated, they undermine trust in digital media. If people can no longer differentiate

between real and fake content, it can erode confidence in online platforms, media outlets, and institutions.

Ethical Guidelines for Addressing Deep Fakes

- 1. Transparency:** AI systems that generate deep fakes should disclose when content is artificial. This transparency helps viewers understand the authenticity of what they see.
- 2. Consent and Ownership:** Individuals should have control over how their likeness is used in AI-generated content. Consent should be obtained before using someone's image or voice in deep fakes.
- 3. Detection and Regulation:** AI-based detection tools can help identify deep fakes, while regulations should be established to deter malicious use. Holding creators accountable for harmful content is essential.
- 4. Education and Awareness:** Educating the public about deep fakes and promoting digital literacy will help people critically assess what they encounter online.
- 5. Ethical Development of AI:** Developers should focus not only on advancing AI technology but also on its ethical implications, ensuring that deep fake technology is used responsibly.

Conclusion

Deep fakes pose complex ethical challenges, as their technology can be both creative and harmful. To mitigate the risks, it is crucial to establish transparency, consent, and accountability in their use. With proper regulation, education, and responsible development, we can ensure that AI-generated content serves the public good rather than undermining privacy and trust.

Conclusion: Core Principles of Ethical AI

As AI continues to play a larger role in shaping our world, understanding and upholding the core ethical principles surrounding its development and application is crucial. Ethics in AI is not a mere afterthought, but a fundamental component that ensures AI serves humanity in responsible and beneficial ways.

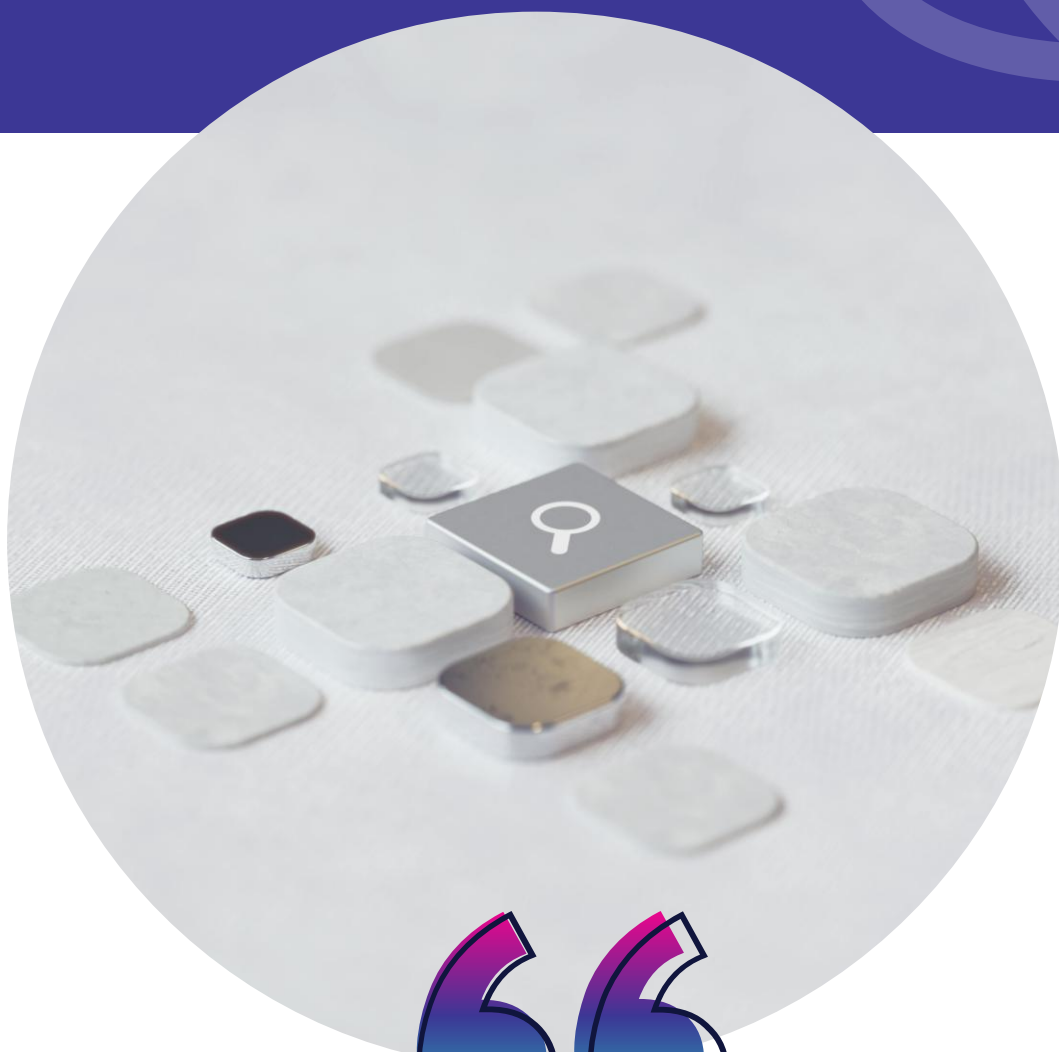
From understanding bias and striving for transparency to ensuring accountability, protecting personal data, and considering the implications of technologies like deep fakes, each principle plays a vital role in shaping the future of AI. These core principles guide developers, policymakers, and users toward creating AI systems that are fair, responsible, and aligned with societal values.

Ultimately, fostering a culture of ethical AI requires collaboration, education, and a commitment to ongoing reflection. As we continue to innovate and integrate AI into various aspects of our lives, it is essential to remember that the ethical decisions we make today will influence the future impact of AI on society. By adhering to these core principles, we can ensure that AI is used to promote equity, transparency, privacy, and accountability, benefiting all members of society.



"The ethical use of AI is not just about doing what is technically possible, but about doing what is right."

Sundar Pichai,
CEO at Alphabet (Google)



03

ETHICAL DILEMMAS IN AI



Ethical Dilemmas of AI

Evaluating Ethical Challenges in AI

As AI technology continues to grow, it introduces increasingly complex ethical challenges. These include not only longstanding concerns like bias and privacy, but also newer issues such as hallucinations by generative AI (e.g. ChatGPT or Google Gemini), over-reliance on flawed input data, and the lack of contextual awareness in decision-making systems. For example, a language model may confidently generate incorrect legal or medical advice, or misinterpret cultural cues in content moderation tasks. Evaluating such risks is critical to ensuring that AI is used responsibly, safely, and fairly. Several tools and frameworks now support developers and educators in navigating the ethical impact of AI systems.

Trustworthy AI

The European Commission's Ethics Guidelines for Trustworthy AI outline key principles for AI systems to be considered lawful, ethical, and robust:

- **Transparency:** AI systems should be explainable to users. For instance, Meta's "Cicero" game-playing AI was praised for combining transparency with natural language strategy—but also highlighted risks of manipulation in language-based AI if systems are not transparent.
- **Accountability:** Developers must be responsible for AI decisions and outcomes. Tools like Model Cards (by Google) and FactSheets (by IBM) provide documentation to clarify AI system purposes and limitations.
- **Privacy and Data Protection:** AI tools like facial recognition must comply with GDPR and national laws. The Clearview AI case—where biometric data was scraped without consent—illustrates how AI can violate privacy rights when safeguards are weak.
- **Non-Discrimination and Fairness:** AI used in recruitment tools (e.g. Amazon's now-shelved hiring algorithm) has shown how training on biased historical data can reinforce gender or

racial discrimination.

- **Safety and Security:** Autonomous systems, like Tesla's Full Self-Driving beta, raise safety concerns when decision-making fails under real-world complexity. Security vulnerabilities in voice assistants like Alexa have also been exploited for unauthorised access.

Ethical AI Tools and Frameworks

Organisations and researchers have created a range of practical tools to help manage AI risks:

1. **Impact Assessments:** AI Impact Assessments—such as the EU's proposed AI Risk Classification under the AI Act—evaluate societal risks before deployment. For example, predictive policing tools have been banned in some European cities after assessments revealed disproportionate harm to minority groups.
2. **Bias Detection Tools:** Tools like IBM's *AI Fairness 360*, Microsoft's *Fairlearn*, and Google's *What-If Tool* enable developers to test AI models for bias. For example, the COMPAS algorithm used in US courts was found to predict recidivism differently by race—highlighting the need for bias audits before real-world use.
3. **AI Governance Frameworks:** The OECD AI Principles and the UNESCO *Recommendation on the Ethics of AI* offer non-binding but influential guidance on trustworthy, human-centred AI. These frameworks help organisations draft internal policies—for instance, requiring human oversight in automated decision-making systems used in education or public services.
4. **Explainability and Interpretability:** Explainable AI (XAI) techniques—such as LIME and SHAP—can reveal which input features most influenced a decision. This is vital in healthcare AI, where black-box models suggesting treatments without rationale (e.g. IBM Watson for Oncology) have come under criticism for lack of clarity and poor real-world accuracy.

Ethical Dilemmas of AI

Evaluating Ethical Challenges in AI

Guiding Principles for Ethical AI Design

To assist in evaluating ethical challenges, developers and AI practitioners can rely on several guiding principles that help frame ethical decision-making:

- **Human-Centered Design:** AI systems should be designed with a focus on the well-being of humans. This means prioritizing user safety, fairness, and respect for human rights in every aspect of AI development.
- **Inclusive Development:** Ethical AI design requires including diverse perspectives, particularly those from marginalized or underrepresented groups, to ensure that AI systems serve everyone equitably.
- **Continuous Monitoring:** Ethical considerations should not be a one-time concern but an ongoing process. Continuous monitoring and auditing of AI systems are necessary to ensure they remain ethical throughout their lifecycle.

Evaluating ethical challenges in AI requires a combination of technical tools, frameworks, and philosophical principles. These tools enable developers, policymakers, and other stakeholders to assess the risks and benefits of AI, ensuring that its use aligns with ethical standards and societal values. As AI continues to evolve, these frameworks will be crucial in guiding its responsible development and application.

READ: The Ethical Considerations of Artificial Intelligence (AI) in Higher Education



Ethical Dilemmas of AI

AI and the Future of Society

Artificial Intelligence is no longer a futuristic concept; it is reshaping the way we live, work, and interact. As AI technology continues to evolve, it brings both exciting possibilities and complex dilemmas that could profoundly impact society. These challenges are particularly important as AI begins to influence critical decisions related to employment, privacy, and even governance. This subchapter explores some of the most significant ethical dilemmas AI presents, including job displacement, surveillance, and AI-driven decision-making.

Job Displacement and the Future of Work

One of the most discussed ethical dilemmas surrounding AI is its potential to displace jobs. AI technologies such as automation, robotics, and machine learning are increasingly capable of performing tasks traditionally carried out by humans. This includes jobs in manufacturing, transportation, customer service, and even fields like healthcare and law. While AI can lead to increased efficiency and productivity, it also raises the question of whether machines will replace human workers, resulting in widespread job loss.

The fear of AI-induced job displacement is not unfounded. For instance, the rise of autonomous vehicles threatens to eliminate jobs for drivers, while automation in warehouses and factories could replace manual labor. According to a study by McKinsey, up to 800 million workers worldwide could be displaced by 2030 due to automation and AI technologies.

However, AI also presents opportunities for job creation in emerging sectors, such as AI ethics, machine learning engineering, and data analysis. The challenge will be ensuring that the workforce is equipped with the skills needed to thrive in a world

where AI plays a dominant role. As AI continues to evolve, education and reskilling programs will be crucial in preparing workers for new roles, reducing the negative impact of job displacement.

Surveillance and Privacy

Another major ethical dilemma posed by AI is the issue of surveillance. AI technologies, such as facial recognition software, data mining, and predictive analytics, have the potential to create an environment of constant monitoring. Governments and corporations are increasingly using AI to track individuals' behaviors, movements, and even emotions, often without their knowledge or consent.

On one hand, AI-driven surveillance can enhance security, help identify criminals, and even prevent potential threats. However, it also raises significant concerns about privacy and the potential for abuse. The mass surveillance of citizens by governments can lead to violations of personal freedoms and human rights. AI technologies are also increasingly used in the private sector, where companies track consumer behaviors for marketing purposes, leading to a more invasive form of surveillance that could erode personal privacy.

The ethical dilemma here is finding a balance between the benefits of AI-driven surveillance and the protection of individual privacy. The need for clear regulations and transparent policies around the use of AI in surveillance is critical. In some countries, regulations such as the General Data Protection Regulation (GDPR) in the European Union have been introduced to limit the extent of surveillance and protect privacy rights. However, more global efforts are needed to ensure that AI surveillance practices do not infringe on fundamental freedoms.

Ethical Dilemmas of AI

AI and the Future of Society

AI in Decision-Making

AI's growing role in decision-making processes brings about another critical ethical dilemma. AI systems are increasingly used in areas like hiring, law enforcement, healthcare, and even judicial systems. These systems analyze vast amounts of data and use algorithms to make decisions that were once made by humans. For example, AI is used to assess job applicants' suitability, predict the likelihood of criminal recidivism, and even diagnose medical conditions.

While AI can process information at a speed and scale that humans cannot, the reliance on algorithms to make decisions also raises significant ethical concerns. One key issue is the potential for bias in AI decision-making. If AI systems are trained on biased data, they can perpetuate and even amplify existing societal inequalities. For example, an AI system used in hiring might favor candidates from certain demographics while overlooking others, resulting in discrimination.

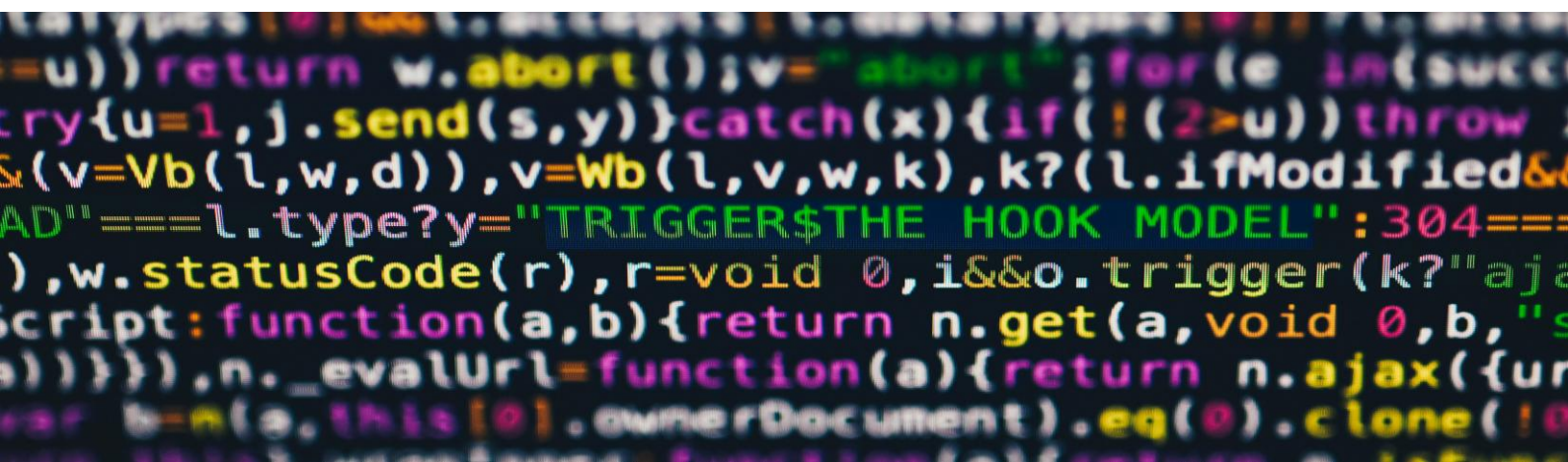
The question of accountability is also critical in AI decision-making. When an AI system makes an incorrect or harmful decision—such as wrongfully denying someone a job or sending an innocent person to jail—who is responsible? Should the

developers, the organizations using the AI, or the AI itself be held accountable? Currently, laws and regulations around AI accountability are still underdeveloped, and clearer frameworks are needed to determine liability in such cases.

AI's Impact on Governance

Finally, AI's potential impact on governance presents another ethical dilemma. With AI's ability to analyze and predict trends based on vast amounts of data, there is the potential for AI to influence political and policy decisions. AI could be used to optimize government operations, improve public services, and create smarter cities. However, the increasing reliance on AI in governance could also lead to power imbalances and unintended consequences.

For instance, AI systems could be used to manipulate public opinion through targeted political ads or fake news, undermining democratic processes. Additionally, if AI systems are used to make policy decisions without sufficient oversight or transparency, it could lead to a loss of public trust in the system. The ethical dilemma here is ensuring that AI is used in governance in ways that promote transparency, fairness, and accountability, while also protecting individual rights and freedoms.



Conclusion: Ethical Dilemmas in AI

As we have explored throughout this chapter, the ethical challenges surrounding AI are vast and complex. From the displacement of jobs and concerns over surveillance to the transparency and accountability of AI systems, these dilemmas touch nearly every aspect of modern life. AI's potential to reshape society offers enormous opportunities but also demands careful consideration of the consequences of its use.

Addressing these ethical challenges requires ongoing dialogue, collaboration, and proactive strategies. The potential for AI to disrupt industries, challenge personal privacy, and influence governance calls for a framework of accountability, fairness, and transparency. As we continue to integrate AI into critical decision-making processes, it is essential to consider the broader social implications, ensuring that the technology aligns with our values and serves the collective good.

Ultimately, the future of AI is not just about technological progress; it is about how we, as a society, choose to navigate these ethical dilemmas. By equipping youth workers, educators, and future generations with the tools and understanding to critically assess AI's impact, we can shape a future where AI is harnessed responsibly and ethically, ensuring its benefits reach all sectors of society without compromising our core values.



"The real question is not whether machines think but whether men do."



B.F. Skinner,
renowned psychologist and
behaviorist.



04

GUIDE YOUTH THROUGH ETHICAL AI DISCUSSIONS



Guiding Youth through Ethical AI Discussions

Facilitating Ethical Discussions

Facilitating discussions around AI ethics can be challenging, especially when introducing complex topics such as bias, privacy, and accountability. However, engaging young people in these discussions is crucial for building their awareness and fostering critical thinking. Here are some approaches to help youth workers and educators facilitate meaningful and productive conversations around the ethical dilemmas in AI:

1. Creating a Safe and Open Environment

For youth to engage with ethical issues surrounding AI, they need a space where they feel comfortable expressing their opinions, asking questions, and challenging ideas. Creating an open and respectful environment helps to ensure that all participants feel safe to share their thoughts without fear of judgment. Encourage active listening and empathy to foster respectful dialogues.

2. Relating AI Ethics to Real-World Examples

Using real-world examples of AI applications can help young people connect abstract ethical concepts to their own lives. Discuss everyday uses of AI, such as social media algorithms, facial recognition technology, and AI-driven job recruitment tools. By framing ethical dilemmas in the context of their own experiences, youth can better understand the impact AI has on their social and professional lives. For instance, explaining how AI algorithms shape what content appears in their feeds can open a conversation about algorithmic bias, privacy concerns, and the responsibility of tech companies.

3. Encouraging Empathy through Role Play and Scenarios

To deepen understanding, use role-playing scenarios where youth can take on different perspectives. For

example, have them role-play as AI developers, users, or policymakers making decisions about how an AI system should be developed or used. This method encourages empathy, as it allows participants to consider the ethical dilemmas from various viewpoints and reflect on how their decisions could impact others. It also helps them understand that AI ethics isn't just about technology, but about human lives and societal well-being.

4. Promoting Open-Ended Questions

Ask open-ended questions that encourage reflection and stimulate deeper thinking. Instead of asking questions with a simple "yes" or "no" answer, encourage young people to explore their reasoning and consider the broader implications of AI decisions. For example:

- “What should happen if an AI system makes a decision that harms someone? Who is responsible?”
- “Do you think AI should be used to monitor people's behaviour in public spaces? Why or why not?”
- “How do we ensure that AI systems are fair and don't discriminate against certain groups of people?”

These types of questions require young people to think critically, consider various factors, and discuss the complexities of ethical AI.



Guiding Youth through Ethical AI Discussions

Facilitating Ethical Discussions

5. Incorporating Different Viewpoints

It is important to expose youth to diverse perspectives, especially since AI's impact spans across various sectors and cultures. Bring in different viewpoints, whether through guest speakers, articles, or videos, from AI experts, ethicists, activists, and technologists from different backgrounds. This helps young people see the multifaceted nature of AI's ethical challenges and consider broader societal implications.

6. Explaining the Concepts of Trustworthy AI

When introducing ethical challenges, it is important to also introduce the concept of "trustworthy AI," which involves AI systems being transparent, accountable, and designed with ethical considerations in mind. Help young people understand how AI can be developed and used responsibly, and emphasize the importance of trust in technology. Use examples from the field of AI ethics, such as Google's AI principles or the European Commission's guidelines for trustworthy AI, to guide discussions on how AI systems can be both innovative and ethical.

7. Encouraging Youth-Led Discussions

Empower young people to lead discussions and generate their own questions about the ethical implications of AI. Provide them with a framework for ethical decision-making, such as principles of fairness, privacy, and transparency, and encourage them to apply these frameworks to their own discussions. This autonomy in decision-making fosters critical thinking and reinforces the idea that they are not just passive recipients of information but active participants in shaping the future of AI.



Guiding Youth through Ethical AI Discussions

Encouraging Critical Thinking

Encouraging critical thinking is one of the most powerful tools educators and youth workers can use to engage young people in discussions about AI ethics. Critical thinking enables individuals to assess the ethical implications of AI technologies, challenge assumptions, and reflect on the long-term consequences of AI applications in society. In this subchapter, we will explore strategies to foster critical thinking among youth and prompt them to question the ethical dimensions of AI.

1. Fostering a Questioning Mindset

One of the key components of critical thinking is the ability to ask questions that challenge the status quo and explore deeper ethical concerns. Encourage young people to adopt a questioning mindset by asking them to consider the “why” and “how” of AI’s impact on society. Instead of simply accepting that AI systems are neutral or that they work as intended, young people should be taught to ask:

- “Who benefits from this AI system, and who might be harmed?”
- “What values are embedded in this AI technology?”
- “How can we ensure this technology respects privacy and fairness?”
- “What are the unintended consequences of using AI in this way?”

These types of probing questions encourage deeper engagement with AI’s societal role and prompt reflection on ethical issues like bias, fairness, and the potential for misuse.

2. Using Case Studies to Challenge Assumptions

Real-life case studies are an excellent way to stimulate critical thinking and engage young people in examining the ethical implications of AI. Present

case studies where AI systems have had unintended or controversial outcomes, such as instances of algorithmic bias, surveillance, or the spread of misinformation. For example:

- **AI in Hiring:** Discuss the case of AI hiring systems that unintentionally favour one gender or ethnic group over others, leading to questions about fairness and bias.
- **Deepfakes in Media:** Present scenarios where deepfake technology has been used to create misleading or harmful content, raising concerns about trust and authenticity in the digital world.
- **Surveillance Systems:** Explore the ethical issues surrounding AI-powered surveillance tools used in public spaces, and discuss the balance between security and personal privacy.

By examining such cases, young people are encouraged to question not only the technology itself but also the underlying assumptions and motivations behind its use. Case studies encourage critical evaluation of how AI systems might perpetuate existing inequalities or create new ethical dilemmas.



WATCH: “AI ethics & engagement with children & young people”

Guiding Youth through Ethical AI Discussions

Interactive Activities for Engagement

Interactive activities are essential tools for engaging young people in ethical AI discussions. By combining theory with hands-on experiences, these activities allow youth to explore AI's ethical dimensions in a more dynamic and engaging way. Practical exercises and role-playing scenarios can help young people critically examine real-world ethical dilemmas, understand diverse perspectives, and develop their ability to engage in thoughtful, informed debates about the implications of AI technologies.

In this subchapter, we will outline several interactive activities designed to stimulate thought, provoke discussion, and help young people actively engage with ethical issues in AI.

1. AI Ethics Debate: The Role of AI in Society

One of the most effective ways to engage young people is through structured debates. This activity allows them to consider opposing viewpoints and defend their positions, helping them to think critically about the ethical issues surrounding AI.

Objective: To explore various ethical dilemmas related to AI and encourage participants to consider both the benefits and challenges associated with AI technologies.

Instructions:

- **Divide the group into two teams:** One team will argue in favour of AI technologies, while the other will argue against them, focusing on ethical concerns such as privacy, bias, and the potential loss of jobs.
- **Research the Topic:** Provide each team with relevant materials and resources on the ethical dilemmas of AI, such as privacy violations, algorithmic bias, and the impact on employment.
- **Set the Debate Rules:** Each team will have a set time to present their arguments, followed by a rebuttal session and a final summary.
- **Guiding Questions:** Use questions like, "Can AI be ethically integrated into public systems?" or "What are the potential risks of AI in decision-making?" to guide the debate.

Outcome: This debate will encourage young people to critically examine AI's ethical implications and sharpen their ability to articulate their views. By engaging with opposing viewpoints, they will learn to evaluate complex ethical issues from multiple perspectives.

Guiding Youth through Ethical AI Discussions

2. AI in the Real World: Case Study Analysis

Case studies provide valuable insights into the real-world applications of AI and its ethical implications. By analyzing actual instances of AI technologies in use, young people can better understand how AI can both benefit society and create ethical dilemmas.

Objective: To allow participants to explore real-life examples of AI applications and discuss their ethical consequences.

Instructions:

- **Present a Case Study:** Choose a case study that highlights an ethical dilemma in AI. For example, you could present the case of a facial recognition system used by law enforcement agencies, or an AI-powered hiring tool that unintentionally discriminates against certain groups of candidates.
- **Group Discussion:** After presenting the case study, divide the group into smaller teams and have them discuss the ethical issues raised. Encourage them to consider questions like:
 - What ethical issues arise from this AI technology?
 - Who is affected by these issues (e.g., individuals, communities, society)?
 - What steps can be taken to address these issues?
- **Present Findings:** Each group should present their analysis, focusing on their proposed solutions to the ethical dilemmas presented in the case study.

Outcome: This activity will help young people develop their analytical skills by evaluating AI technologies through an ethical lens. It will also encourage them to consider how they might address AI's ethical challenges in practical, actionable ways.



**READ: Artificial Intelligence -
examples of ethical dilemmas**

Guiding Youth through Ethical AI Discussions

3. Role-Playing: Who is Responsible?

Role-playing activities can help young people empathize with various stakeholders involved in AI-related ethical dilemmas. This exercise helps participants consider how different individuals or organizations might approach ethical issues and the impact of their decisions on others.

Objective: To engage youth in exploring accountability and responsibility in AI systems.

Instructions:

- **Assign Roles:** Provide participants with different roles to play in a scenario where an AI system causes an ethical dilemma. Examples of roles include:
 - The developer who created the AI system.
 - The government regulator responsible for overseeing AI use.
 - A consumer who is impacted by the AI's decisions.
 - An activist advocating for the responsible use of AI.
- **Scenario Setup:** Present a scenario, such as a biased AI system used for loan approvals that disproportionately denies credit to people from certain communities. The participants will debate who is responsible for the harm caused by the AI system.
- **Discussion and Decision-Making:** Each role-holder must argue their position and suggest actions to remedy the ethical issue. After a set time, allow the participants to discuss and vote on the best course of action to address the dilemma.

Outcome: This role-playing activity will help youth consider the perspectives of multiple stakeholders and think critically about accountability in AI systems. It will encourage them to reflect on the complex nature of responsibility in the digital age and how various parties contribute to ethical decision-making.

**READ: Role Playing | Center
for Innovative Teaching and
Learning**



Guiding Youth through Ethical AI Discussions

4. Design Your Ethical AI

This creative activity allows young people to think through the process of designing an AI system from scratch while considering its ethical implications. By engaging in this hands-on activity, participants will gain a better understanding of how ethical principles can be embedded into the development process.

Objective: To encourage youth to think about how AI systems can be designed in ways that align with ethical principles.

Instructions:

- **Team Formation:** Divide participants into small teams and give each team the task of designing an AI system. They could choose a specific application, such as an AI for hiring, healthcare diagnostics, or smart city management.
- **Ethical Design:** Each team must consider and integrate ethical principles into their design. They should discuss and address issues such as fairness, transparency, data privacy, and inclusivity.
- **Presentation:** After designing their AI system, each team will present their system and explain how they incorporated ethical principles into the design. They should also discuss the potential ethical challenges their AI system might face and how they would address them.

Outcome: This activity fosters creativity and ethical thinking, allowing young people to actively participate in the process of designing AI systems that prioritize ethical considerations. It helps them understand that ethical AI design is not an afterthought but a critical aspect of the development process.



READ: Artificial Intelligence - examples of ethical dilemmas

Guiding Youth through Ethical AI Discussions

5. AI Ethics Quiz

A fun and interactive way to test and reinforce what young people have learned about AI ethics is through a quiz. This can be done as an individual or team activity, encouraging youth to reflect on key concepts in a lighthearted but educational way.

Objective: To assess young people's understanding of AI ethics in a fun and engaging format.

Instructions:

- **Prepare Questions:** Create a series of multiple-choice or short-answer questions based on the content covered in the discussions and activities. Questions could cover topics such as:
 - What is algorithmic bias?
 - Why is transparency important in AI systems?
 - What are the potential risks of AI in decision-making?
 - Who is responsible when AI systems cause harm?
- **Conduct the Quiz:** Ask the questions aloud or display them on a screen. Participants can answer individually or as teams, and the correct answers should be explained after each question.

Outcome: This quiz helps reinforce the learning process and provides an opportunity for youth to test their knowledge. It's a fun way to encourage reflection and solidify understanding of key ethical AI concepts.

READ: Interactive AI Ethics Quiz



How to Conduct Classes on AI and Ethics

A quick, practical checklist for educators and youth workers

1

Prepare the Groundwork

- ☐ Familiarise yourself with key ethical principles (e.g. transparency, fairness, accountability, data protection).
- ☐ Review relevant frameworks (EU AI Act, GDPR, UNESCO Ethics of AI).
- ☐ Assess your group's digital literacy and prior knowledge of AI.
- ☐ Choose case studies that are age-appropriate and culturally relevant.

2

Design an Inclusive and Engaging Session

- ☐ Use real-life examples (e.g. biased recruitment tools, facial recognition, hallucinations in chatbots).
- ☐ Balance technical concepts with social impact discussions.
- ☐ Encourage critical thinking through questions like:
 "Should AI decide who gets a job?"
 "Is it ethical to use AI in education or policing?"
- ☐ Include different viewpoints and promote open dialogue.

3

Use Interactive Learning Methods

- ☐ Incorporate multimedia (videos, infographics, simulations).
- ☐ Use interactive tools (e.g. AI Fairness 360 demos, Google's What-If Tool).
- ☐ Run ethical dilemmas or roleplay debates ("You are a developer... what do you do?").
- ☐ Use a short quiz or poll to spark reflection and engagement.

4

Focus on Values and Responsibility

- ☐ Link lessons to digital citizenship, equality, and sustainability.
- ☐ Discuss the human choices behind AI—data selection, design decisions, intended outcomes.
- ☐ Emphasise that "just because we can doesn't mean we should."
- ☐ Highlight the importance of diverse, inclusive tech teams to reduce bias.

5

Connect to Daily Life and Future Careers

- ☐ Show how AI affects daily tools (e.g. TikTok algorithms, voice assistants, auto-translation).
- ☐ Explore ethical roles in AI development (AI policy advisor, ethical UX designer, etc.).
- ☐ Invite students to identify ethical risks in apps or AI tools they use.
- ☐ Encourage curiosity: "What problems would you solve with ethical AI?"

6

Evaluate and Reflect

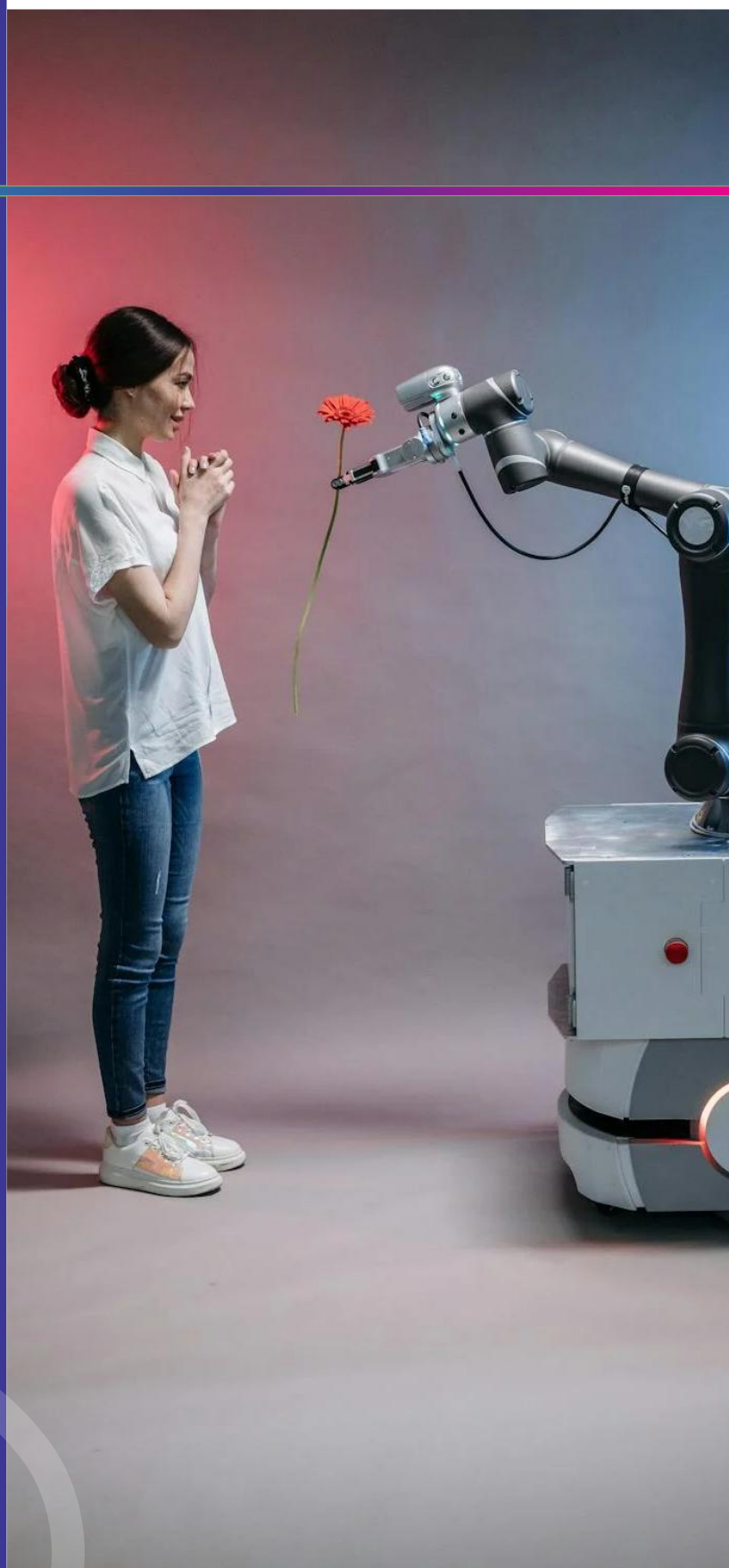
- ☐ Use short feedback forms or reflection exercises.
- ☐ Ask: "What was the most surprising thing you learned?"
- ☐ Adjust future sessions based on engagement, understanding, and feedback.

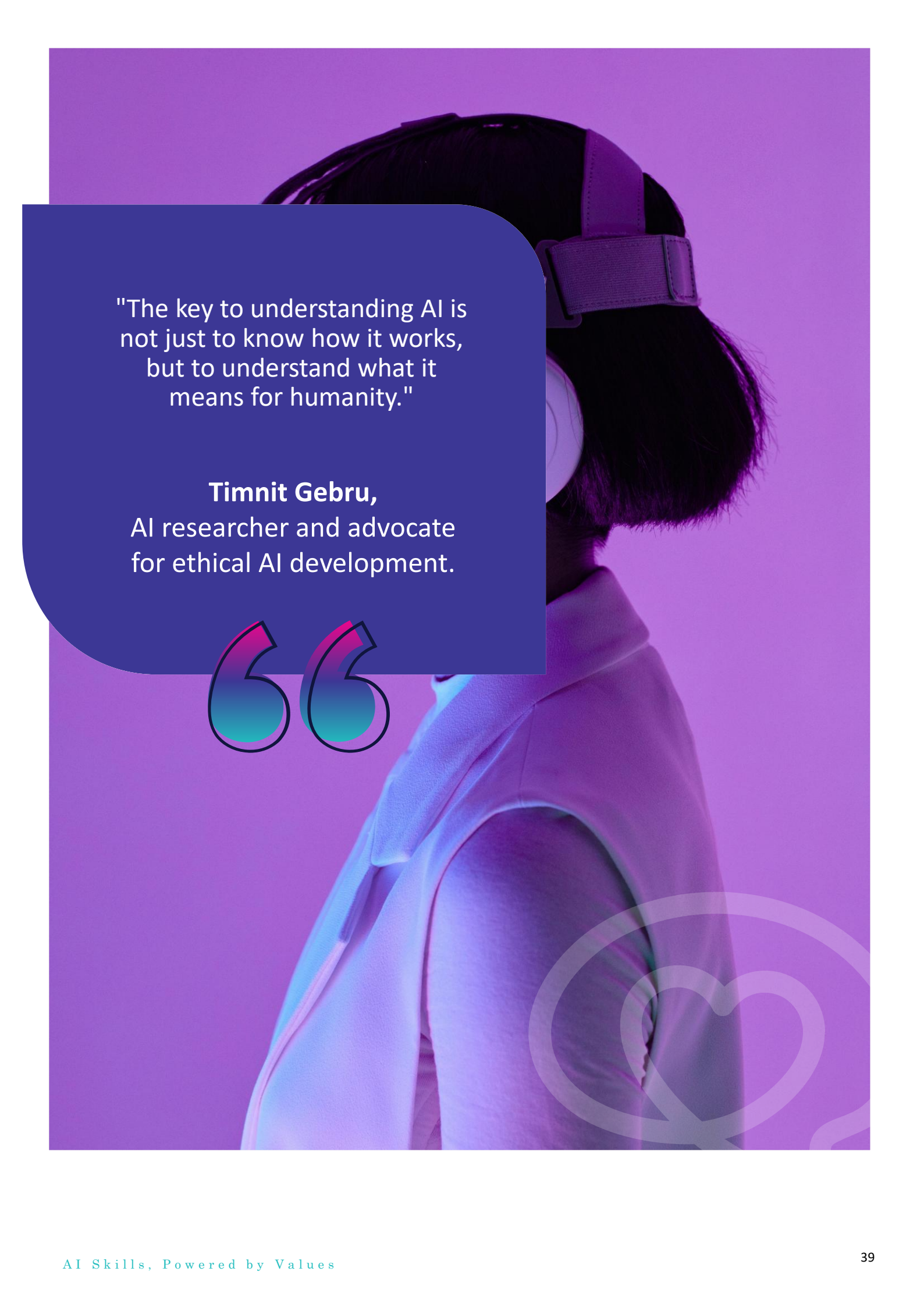
Conclusion: Guiding Youth through Ethical AI Discussions

In this chapter, we explored the importance of guiding young people through the complex ethical considerations surrounding AI. By facilitating discussions, encouraging critical thinking, and incorporating interactive activities, youth can be empowered to understand not only how AI works but also its societal impact and ethical implications. These tools help foster a sense of responsibility, enabling young people to engage thoughtfully with AI technologies and their role in shaping the future.

The activities outlined in this chapter—ranging from debates and role-playing to case studies and creative design exercises—provide practical ways for educators and youth workers to inspire informed, open dialogue about AI ethics. They serve as both educational and transformative experiences that challenge young people to critically examine the technology they use and encounter in their daily lives.

By helping youth navigate these ethical dilemmas, we equip them with the skills to make responsible decisions and contribute to discussions about the role AI will play in shaping their lives, careers, and society at large. This chapter is a call to action for educators and youth workers to become facilitators of these important conversations, ultimately guiding youth to become ethically aware and active participants in the AI-driven world of tomorrow.





"The key to understanding AI is not just to know how it works, but to understand what it means for humanity."

Timnit Gebru,
AI researcher and advocate
for ethical AI development.



05

ANNEXES



Annexes

References

1. **Binns, R. (2018).** *"On the Logic of Trustworthy AI."* Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Retrieved from: <https://dl.acm.org/doi/10.1145/3173574.3173793>
2. **Bryson, J. (2018).** *"The Ethics of Artificial Intelligence: An Overview."* Oxford Handbook of Ethics of AI. Oxford University Press.
3. **Dastin, J. (2018).** *"Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women."* Reuters. Retrieved from: <https://www.reuters.com/article/us-amazon-com-jobs-insight-idUSKCN1MK08G>
4. **European Commission (2020).** *"White Paper on Artificial Intelligence – A European Approach to Excellence and Trust."* European Commission. Retrieved from: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
5. **Floridi, L. (2018).** *"The Ethics of Artificial Intelligence."* The Cambridge Handbook of Information and Computer Ethics. Cambridge University Press.
6. **Gebru, T. (2020).** *"AI and the Future of Humanity."* The New York Times. Retrieved from: <https://www.nytimes.com/2020/09/03/technology/ai-debate-ethical.html>
7. **Gillespie, T. (2018).** *"The Relevance of Algorithms in Social Media."* Media, Culture, & Society. Retrieved from: <https://journals.sagepub.com/doi/10.1177/0163443718768324>
8. **Goodall, J. (2019).** *"Ethics and Accountability in AI Systems."* Stanford Encyclopedia of Philosophy. Retrieved from: <https://plato.stanford.edu/entries/ethics-ai/>
9. **O'Neil, C. (2016).** *"Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy."* Crown Publishing Group.
10. **Raji, I. D., & Buolamwini, J. (2019).** *"Actionable Auditing: Investigating the Impact of Publicly Available Face Dataset on Racialized Demographic Groups."* Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Retrieved from: <https://dl.acm.org/doi/10.1145/3293663.3293674>
11. **Russell, S., & Norvig, P. (2016).** *"Artificial Intelligence: A Modern Approach."* 3rd Edition. Pearson.
12. **Simonite, T. (2018).** *"How Deep Learning Algorithms Are Tackling AI's Bias Problem."* MIT Technology Review. Retrieved from: <https://www.technologyreview.com/2018/02/26/65985/how-deep-learning-algorithms-are-tackling-ais-bias-problem/>
13. **Vincent, J. (2021).** *"What Are Deep Fakes and How Do They Work?"* The Verge. Retrieved from: <https://www.theverge.com/2021/1/20/22242362/deepfake-ai-fake-video-deep-learning-explained>
14. **Zeng, J. (2020).** *"Transparency in AI and Ethics: The Importance of Open AI Systems."* Journal of Ethics and Information Technology. Retrieved from: <https://link.springer.com/article/10.1007/s10676-020-09528-3>

Annexes

Glossary

Artificial Intelligence (AI)

A branch of computer science that deals with creating machines or software capable of performing tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and language translation.

Algorithm

A set of rules or instructions followed by a computer to solve a problem or complete a task. In AI, algorithms help machines process data and make decisions.

Bias in AI

The presence of systematic and unfair discrimination in AI systems due to imbalanced data or flawed algorithmic design, leading to unfair decisions for certain groups.

Data Privacy

The practice of protecting personal information and ensuring that it is only used in ways that the individual has agreed to. In AI, data privacy refers to how personal data is collected, stored, and used by AI systems.

Deep Learning

A subset of machine learning that involves neural networks with many layers. Deep learning models are particularly effective for tasks such as image and speech recognition.

Ethical AI

AI systems that are designed, developed, and deployed in a manner that aligns with ethical principles such as fairness, accountability, transparency, and respect for privacy.

Machine Learning (ML)

A type of AI where algorithms learn from data and improve their performance over time without

explicit programming. It is used for tasks like predictive analysis, pattern recognition, and anomaly detection.

Natural Language Processing (NLP)

A field of AI focused on enabling machines to understand, interpret, and respond to human language. It is the foundation for technologies like speech recognition and chatbots.

Transparency in AI

The principle that AI systems should be open and understandable, with clear explanations of how decisions are made and the ability to trace and audit the processes behind AI outcomes.

Accountability in AI

The concept that the creators and users of AI systems should be responsible for the outcomes of those systems. If AI makes an error or causes harm, it is important to know who is responsible.

Civic Engagement

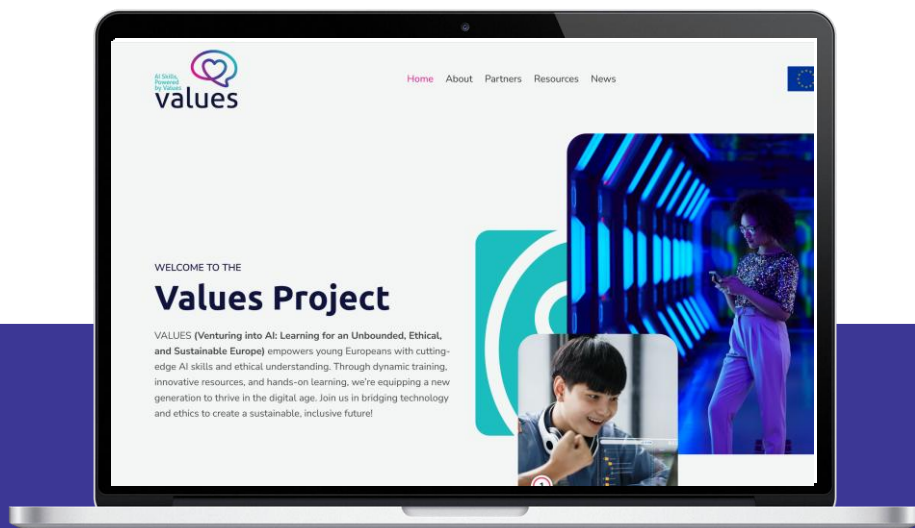
The involvement of individuals in the community, including participation in decision-making processes and using technologies (such as AI) to contribute to society in a responsible manner.

Fairness in AI

Ensuring that AI systems treat all individuals and groups equitably, without favoritism or discrimination. It involves ensuring that biases in the data or the model do not lead to unfair outcomes.

Responsible AI

The development and deployment of AI systems that ensure safety, fairness, and inclusivity. It includes ensuring that AI applications do not cause harm to individuals or society and that they uphold ethical standards.



www.valuesai.eu

follow our journey



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Ethical AI Fundamentals Guide © 2025 by VALUES is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by-sa/4.0/)



Starter Kit for Ethical AI

Practical Applications of AI Guide

www.valuesai.eu



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Practical Applications of AI Guide © 2025 by VALUES is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by-sa/4.0/)

Introducing the VALUES Project

The **VALUES project** is a forward-thinking initiative aimed at empowering young people across Europe by providing them with the essential skills and ethical understanding needed to navigate the rapidly developing world of artificial intelligence (AI). As AI technologies become an integral part of everyday life, it is essential for youth to not only acquire technical knowledge but also to develop a strong awareness of the ethical considerations surrounding AI. The project's core objective is to equip young people, particularly those facing employment challenges, with the knowledge and tools to critically evaluate AI technologies, ensuring they can engage with these systems responsibly. By fostering a deeper understanding of AI ethics, the VALUES project aims to prepare young people to become responsible digital citizens and ethical innovators, ready to contribute positively to the digital economy. The anticipated outcomes include increased AI literacy, a more ethical approach to technology use, and enhanced employment opportunities in the growing AI sector.

What is the VALUES Starter Kit for Ethical AI?

*The **VALUES Starter Kit for Ethical AI** is your gateway to exploring the fascinating and essential world of ethical AI. Designed for youth workers and educators, this dynamic resource offers you the tools to dive into the complexities of AI, from its core principles to the ethical dilemmas it presents. Packed with real-world examples, practical strategies, and*

engaging activities, this kit will help you spark meaningful discussions with young people about the impact of AI on their lives. Whether you're introducing AI for the first time or guiding young people through its ethical implications, this Starter Kit will empower you to teach with confidence and ignite curiosity about the responsible use of AI in today's world.

contents

01	INTRODUCTION TO AI IN YOUTH WORK
02	AI-ASSISTED LESSON PLANNING AND DEVELOPMENT
03	AI FOR CREATIVITY AND DESIGN
04	AI FOR INCLUSION AND ACCESSIBILITY IN EDUCATION
05	AI FOR MONITORING AND PERSONALISED LEARNING
06	ETHICAL AI USAGE AND CRITICAL THINKING
07	CONCLUSION: PREPARING YOUTH FOR AN AI FUTURE
08	ANNEXES



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.
Practical Applications of AI Guide © 2025 by VALUES is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by-sa/4.0/)

01

INTRODUCTION TO AI IN YOUTH WORK



Introduction to AI in Youth Work

Overview of AI's Role in Education and Youth Work

Artificial Intelligence (AI) is rapidly transforming how we work, live, and learn, and its potential in education is vast. From creating personalised learning experiences to providing real-time insights into student performance, AI is reshaping traditional educational practices. As youth workers, educators, and trainers, it's essential to understand AI's role and how it can enhance the youth work environment.

AI in education refers to the use of algorithms, data, and automation to improve learning outcomes, streamline administrative tasks, and foster innovation. In youth work, AI tools can offer personalised learning experiences, support accessibility, and promote creativity. These benefits help youth workers focus on the core aspects of their work—guiding, mentoring, and empowering young people—while AI takes care of time-consuming administrative tasks.

Some AI tools allow for adaptive learning systems, where the content and pace adjust based on the learner's progress. This ensures that every individual, whether struggling or excelling, receives a learning experience tailored to their needs. For youth workers, AI can be an invaluable resource in developing lesson plans, tracking student progress, and identifying areas where support is needed.

In addition to personalising learning, AI also opens doors to innovative, creative possibilities. AI tools can help youth express themselves through art, music, and video, providing a new medium for creative exploration. With the right tools, youth workers can engage young people in exploring the intersection of technology and creativity, encouraging problem-solving, collaboration, and critical thinking.

Furthermore, AI has the potential to make education more inclusive. Tools designed for accessibility can assist youth with learning disabilities, language barriers, and other challenges, ensuring that all students have equal opportunities to succeed. For instance, AI can provide real-time translation for non-native speakers or read aloud text for students with dyslexia, helping them engage with the material in ways that work for them.

However, the integration of AI into youth work and education also raises important ethical considerations. As we introduce more AI-driven tools into classrooms and training sessions, it's crucial to instil a sense of responsibility in young people. They should not only learn how to use AI but also understand its ethical implications. Discussions around data privacy, algorithmic bias, and fairness should be part of the educational journey, helping young people develop critical thinking skills and become responsible digital citizens.

AI's role in education and youth work is evolving, and as technology advances, so too will its impact. By integrating AI thoughtfully and ethically, youth workers can enhance learning experiences, promote creativity, and ensure inclusivity, all while preparing young people for a future in a technology-driven world.

This chapter serves as an introduction to how AI can be utilised effectively in youth work and education. Throughout this guide, we'll explore specific AI tools that youth workers can implement in their day-to-day activities, making lesson planning more efficient, promoting creativity, enhancing accessibility, and fostering ethical conversations around AI.

Introduction to AI in Youth Work

Importance of Integrating AI into Daily Practices

Integrating Artificial Intelligence (AI) into the daily practices of youth work and education isn't just about adopting new technology—it's about enhancing the quality of education, making learning more personalised, efficient, and accessible, and preparing young people for the future. In a world where AI is increasingly present in all sectors, from healthcare to business, youth workers must embrace these tools to stay relevant and to equip young people with the skills they need to succeed in a rapidly evolving society.

Enhancing Personalisation and Engagement

One of the greatest benefits of integrating AI into daily practices is the ability to offer personalised learning experiences. AI can adapt content and teaching methods to meet the specific needs of each learner, based on their individual strengths and weaknesses. For example, AI can analyse student responses to assess their knowledge gaps, automatically adjusting the difficulty of the material and offering additional support where needed. This ensures that all learners, whether they're struggling or excelling, can engage with the material at an appropriate level.

Youth workers can use AI tools to create custom lesson plans and activities that resonate with the individual interests and needs of their learners. Whether it's adjusting the pace of a lesson for a slower learner or challenging more advanced students with deeper content, AI ensures that no one is left behind, fostering a more inclusive learning environment.

Time-Saving and Efficiency

AI can significantly reduce the time spent on administrative tasks, allowing youth workers to focus more on their interactions with young people. For example, AI tools can automate the creation of

quizzes, summaries, and lesson plans, saving valuable time. Platforms like [Google Classroom](#) and AI-powered quiz generators such as [Quillionz](#) help youth workers rapidly generate resources and track student performance, allowing for quicker adjustments to teaching strategies and ensuring lessons remain dynamic and relevant.

Moreover, AI can automate attendance tracking, data entry, and grading, reducing the administrative burden that often weighs on youth workers. This increased efficiency can free up more time for one-on-one mentoring, youth engagement, and problem-solving, which are essential components of effective youth work.

Fostering Critical Thinking and Creativity

In addition to personalisation, AI can inspire creativity and critical thinking in young people. AI tools like [Runway ML](#) and [DeepDream Generator](#) allow students to experiment with creative projects such as artwork, music, and video creation, all with the support of AI. These tools can help youth discover their creative potential while gaining exposure to cutting-edge technology. By using AI to explore artistic expression, young people are able to blend technology with creativity, preparing them for future careers in industries like digital media, design, and technology.

Youth workers can also use AI to foster critical thinking about technology itself. Introducing young people to AI not only gives them the opportunity to learn how to use it but also enables them to consider its ethical implications, such as algorithmic bias and privacy issues. By involving young people in conversations about the responsible use of AI, youth workers can help them become not just consumers of technology, but informed, ethical creators and users of it.

Introduction to AI in Youth Work

Importance of Integrating AI into Daily Practices

Encouraging Inclusivity and Equal Opportunities

AI has the potential to address longstanding barriers in education by making learning more accessible to everyone, including those with disabilities. Tools like [Microsoft's Immersive Reader](#) support students with dyslexia, while [Otter.ai](#) provides live transcriptions, helping students with hearing impairments fully participate in lessons. By integrating these AI tools into daily practices, youth workers can create a more inclusive environment, ensuring that all young people, regardless of their background or abilities, have an equal opportunity to learn and succeed.

For youth workers, the importance of integrating AI into daily practices also lies in preparing young people for the future workforce. As AI continues to shape industries across the globe, understanding and working with AI will be a fundamental skill. Equipping young people with AI literacy—whether through hands-on experiences with AI tools or discussions about its societal impact—ensures they are not only ready for the careers of today but also the opportunities of tomorrow.

Ethical Considerations and Responsible Use

As youth workers integrate AI into their daily practices, it's crucial to consider the ethical implications of using such technology. AI can bring about positive change, but it also raises concerns about data privacy, bias, and fairness. Therefore, it's essential to not only teach young people how to use AI but also how to critically evaluate its ethical dimensions. Introducing AI in a responsible and thoughtful way ensures that young people understand the impact technology has on society and equips them with the tools to make informed, ethical decisions as they navigate a tech-driven world.

By integrating AI into their daily practices, youth workers can significantly enhance the learning experience, make education more inclusive, and help young people develop the skills they need to thrive in an AI-powered future. The integration of AI tools empowers youth workers to innovate in the classroom, allowing them to focus on their core mission: mentoring, guiding, and fostering the development of young people.

READ: Enhancing Youth Work through Artificial Intelligence



Introduction to AI in Youth Work

How AI Can Enhance Engagement, Creativity, and Inclusivity in Learning

AI has the potential to significantly enhance the learning experience by making it more engaging, creative, and inclusive. By leveraging AI, youth workers can cater to diverse learning styles, unlock new avenues for creativity, and ensure that no young person is left behind in the learning process. Here's how AI can contribute to each of these areas:

Enhancing Engagement in Learning

Engagement is a cornerstone of effective learning. AI-driven tools can make learning more interactive and dynamic, offering a level of personalisation that traditional methods simply cannot match.

1. Personalised Learning Experiences

AI can tailor educational content to meet the individual needs of learners, creating a more engaging experience. For example, platforms like [DreamBox Learning](#) use AI to analyse student responses in real-time and adjust the level of difficulty to match their pace, keeping learners engaged by ensuring they are constantly challenged at a level suited to their current abilities. This dynamic adaptation of content keeps students motivated and immersed in their learning journey, as they are neither overwhelmed nor bored.

2. Gamification and Interactive Tools

AI can also enhance engagement through gamification, where learners earn points, badges, or rewards for completing tasks. [Classcraft](#), an AI-powered platform, uses game mechanics to engage students in learning, tracking their progress, and rewarding them for their involvement in class activities. Such tools motivate young people to take an active role in their education by turning learning into a fun, goal-oriented experience.

3. AI in Social and Collaborative Learning

AI can facilitate collaboration by connecting students

with peers around the world. Tools like **Google Classroom**, integrated with AI, not only help youth workers manage and distribute lessons but also encourage collaborative work through online discussions and group projects. AI tools can monitor group interactions, provide insights into individual contributions, and suggest ways to improve teamwork, fostering a deeper sense of engagement in the learning process.

Unlocking Creativity through AI

AI opens up new possibilities for creative expression, allowing young people to explore and experiment in ways that were previously unimaginable.

1. AI as a Creative Partner

AI tools like [Runway ML](#) allow young people to create artwork, animations, and videos by simply providing input and letting the AI generate visually striking content. These tools remove barriers to entry for young people who may not have traditional artistic skills but are interested in creative expression. Whether they are creating digital art, producing short films, or experimenting with interactive installations, AI serves as a collaborative partner, enhancing their creative abilities.

2. AI-Driven Storytelling

AI can also assist young people in telling their stories. Tools like [DeepDream Generator](#) allow users to create abstract, AI-generated visuals based on their input, encouraging creative exploration. These images can serve as the basis for storytelling in various formats, from written stories to video projects. AI's ability to process vast amounts of data and turn it into artistic content opens doors for youth to explore new forms of creativity, including digital art, music, and narrative storytelling.

Introduction to AI in Youth Work

How AI Can Enhance Engagement, Creativity, and Inclusivity in Learning

3. Innovation in Music and Performing Arts

AI is also finding its place in music and the performing arts. AI-driven software like [Amper Music](#) and [AIVA](#) helps young people compose music by generating unique soundtracks based on specific genres, tempos, or emotional tones. For aspiring musicians or performers, these tools provide a springboard for experimentation, offering new ways to create and produce music that they may not have otherwise imagined. This creativity-driven approach helps young people think outside the box and explore AI's artistic potential.

Fostering Inclusivity in Education

AI plays a crucial role in making education more inclusive by addressing the diverse needs of learners, ensuring that every young person, regardless of ability, has access to quality education.

1. AI for Learners with Disabilities

AI can help youth workers ensure that all students have access to the learning materials and support they need, particularly those with disabilities. For instance, [Microsoft Immersive Reader](#) is an AI-powered tool that helps learners with dyslexia or visual impairments by reading text aloud and offering translation and text-to-speech capabilities. These tools ensure that students with different learning needs can access and engage with content on an equal footing with their peers.

2. Real-Time Transcriptions and Support for Hearing-Impaired Students

For students with hearing impairments, AI tools like [Otter.ai](#) provide real-time transcriptions of spoken words, turning classroom discussions into text that can be read on a screen. This inclusion of real-time transcription ensures that hearing-impaired students

can follow lessons and participate in discussions, enhancing their engagement and understanding of the content.

3. Supporting Students with Learning Difficulties

AI tools can also provide extra support for students with learning difficulties. [Read&Write](#), an AI tool for students with learning disabilities, helps them by offering features like word prediction, speech-to-text, and screen reading. These tools ensure that all students, regardless of their abilities, can engage in the same educational opportunities. AI's role in supporting diverse learners in a flexible and tailored manner helps foster an inclusive learning environment where every young person can thrive.

Breaking Down Barriers to Learning

AI helps break down traditional barriers to education by providing new ways to access knowledge, whether through self-paced learning, interactive tools, or real-time feedback. AI's potential to assist in remote learning, provide continuous support, and ensure educational opportunities for underserved or disadvantaged groups is especially relevant in today's global landscape, where access to education can often be limited by geographical or socio-economic factors.

By integrating AI tools into everyday educational practices, youth workers can create an environment where learning is personalised, creative, and accessible. These advancements enable young people to engage in their education in a way that is tailored to their individual needs, interests, and abilities, ultimately promoting greater equity in education and preparing them for the challenges and opportunities of an AI-driven future.

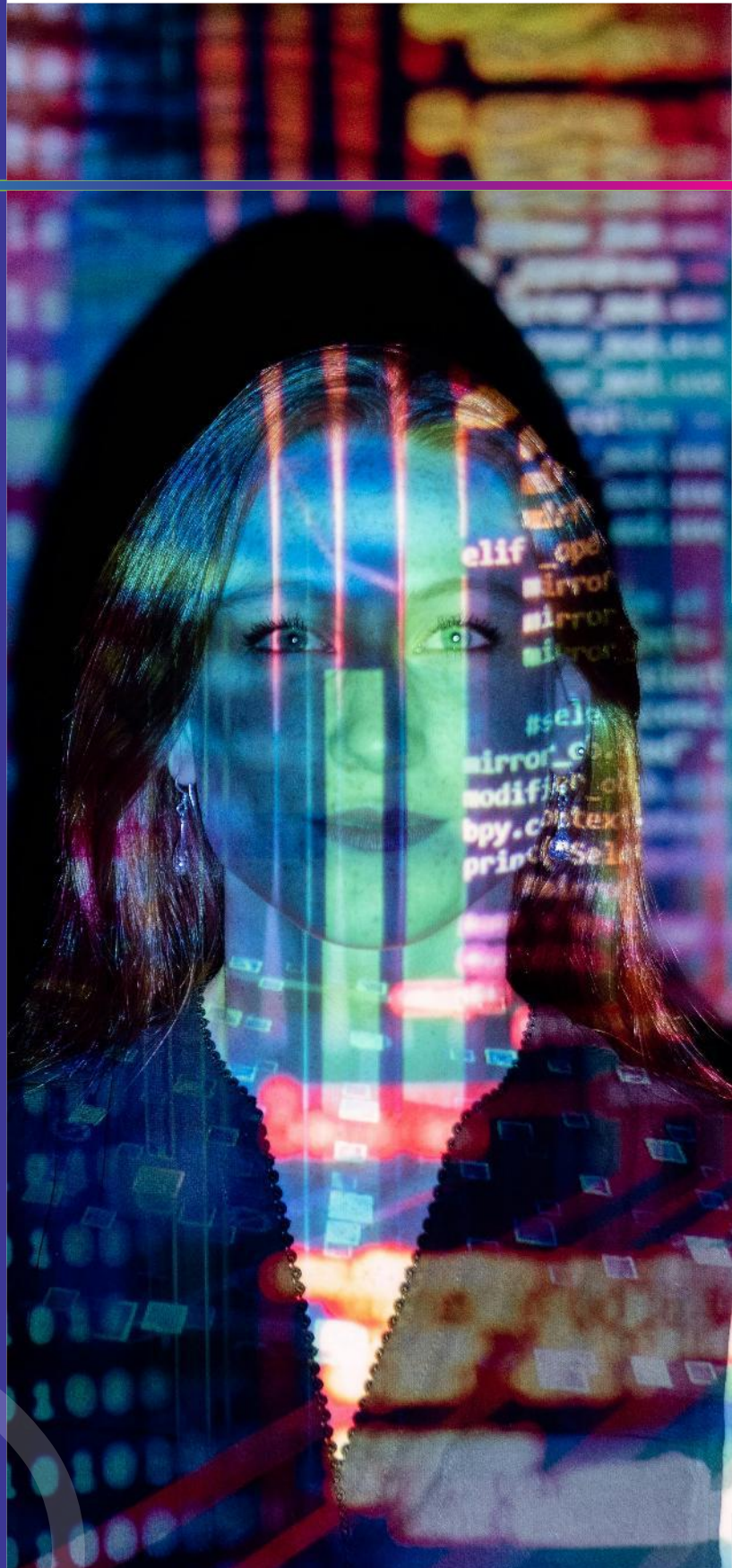
Conclusion:

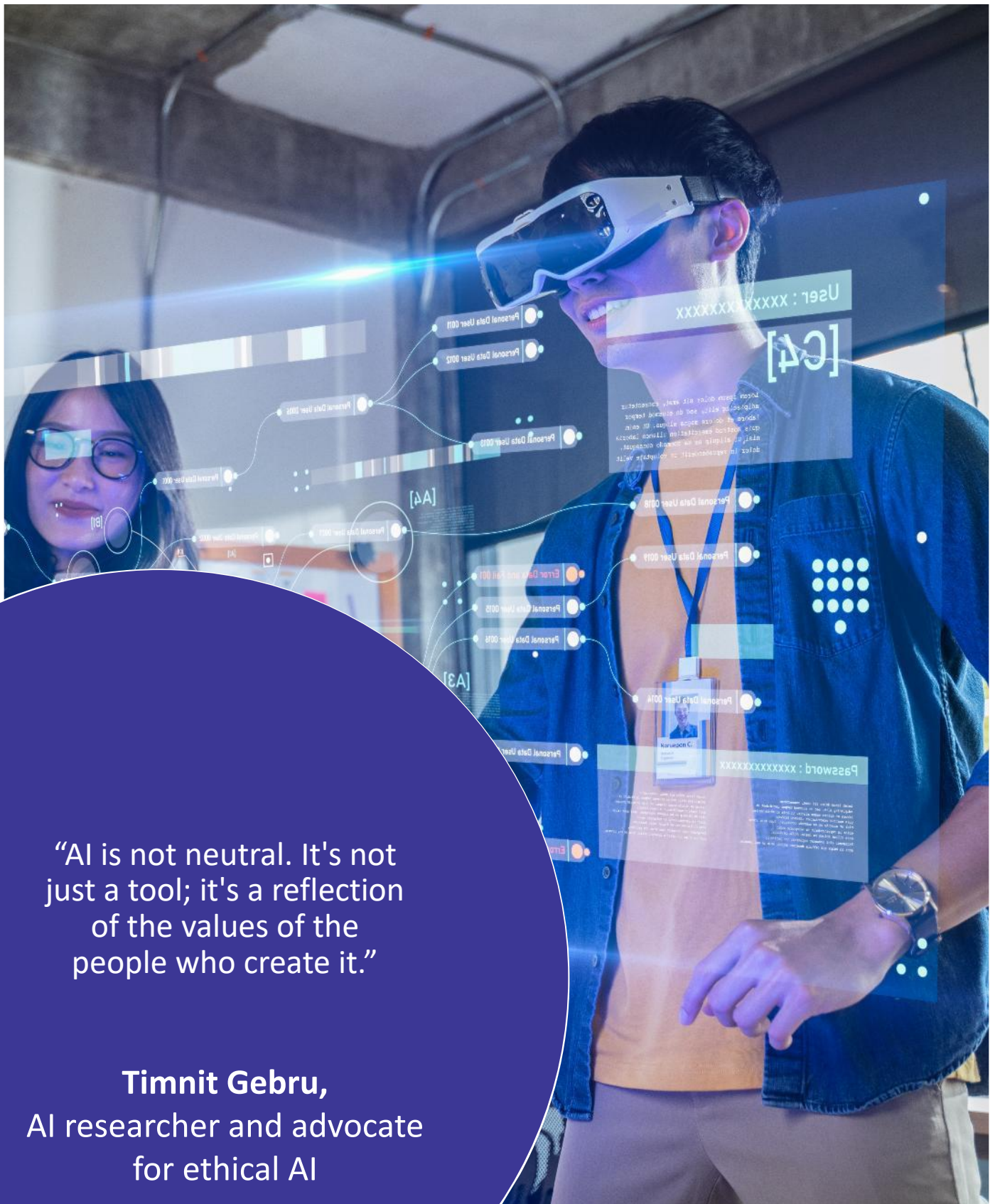
Introduction to AI in Youth Work

Incorporating AI into youth work and education is not just a technological advancement—it is a powerful tool that can transform the way young people learn, engage, and grow. As we've explored in this chapter, AI has the potential to enhance engagement, foster creativity, and promote inclusivity in learning environments. By personalising educational content, offering new creative possibilities, and ensuring that all learners, regardless of their background or abilities, have access to the tools they need, AI empowers both youth workers and the young people they support.

The importance of integrating AI into daily practices cannot be overstated. AI has the power to redefine the educational experience, providing youth workers with the resources to meet the diverse needs of their learners and to facilitate deeper, more meaningful interactions. Through its ability to adapt and personalise learning pathways, AI can make education more dynamic, relevant, and inclusive.

As we continue to explore the practical applications of AI throughout this guide, it becomes clear that AI is not just a fleeting trend—it is here to stay. Its integration into youth work offers new opportunities to improve learning outcomes, boost engagement, and ensure a more equitable educational experience for all young people. The future of youth work, with AI as a partner, holds exciting possibilities for young people to develop the skills and mindset needed for the challenges and opportunities of an increasingly AI-driven world.





“AI is not neutral. It's not just a tool; it's a reflection of the values of the people who create it.”

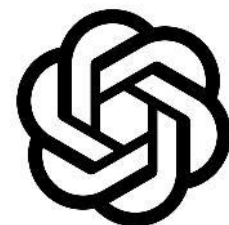
Timnit Gebru,
AI researcher and advocate
for ethical AI



02

AI-ASSISTED LESSON PLANNING AND DEVELOPMENT





AI-Assisted Lesson Planning and Development

ChatGPT (OpenAI)

<https://openai.com/chatgpt/overview/>

ChatGPT, developed by OpenAI, is a powerful AI tool that can significantly streamline lesson planning for youth workers. With its ability to understand natural language, ChatGPT can generate customised lesson plans, quizzes, and discussion prompts based on specific topics and educational goals. This allows youth workers to quickly create relevant, engaging, and personalised learning content.

Customised Lesson Plans

ChatGPT can help youth workers create detailed lesson plans tailored to particular topics, age groups, and learning objectives. By inputting basic instructions, such as the topic (e.g., climate change) or the desired learning outcomes, youth workers receive a structured plan with key concepts, activities, and discussion points. This helps reduce planning time while ensuring lessons are relevant and accessible to all learners, no matter their needs or learning styles.

Quizzes and Assessments

Creating quizzes can also be simplified using ChatGPT. By requesting questions based on the lesson material, youth workers can receive multiple-choice, short-answer, or true/false questions tailored to the content. ChatGPT can even generate answer keys, providing a quick way to assess students' understanding and offer real-time feedback.

Discussion Prompts

Engaging students in meaningful conversations is made easier with ChatGPT's ability to generate thought-provoking discussion prompts. These prompts can cover a range of topics, such as ethical dilemmas or current events, encouraging critical

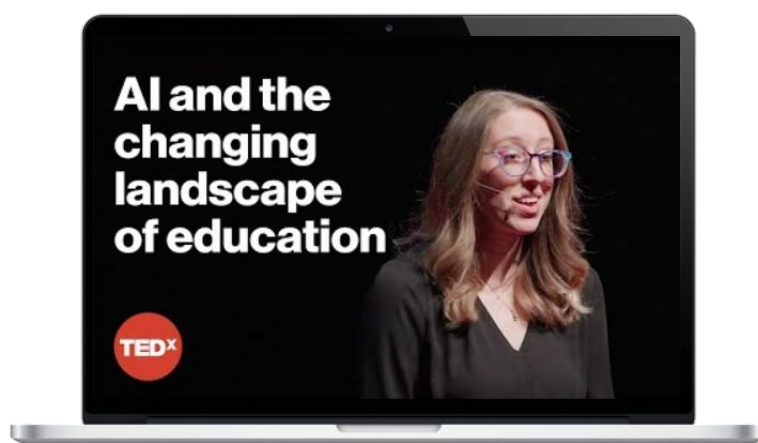
thinking and fostering deeper conversations among youth. For example, ChatGPT might suggest questions like, "What are the ethical implications of AI?" or "How do you think AI will impact future jobs?" These prompts help stimulate reflection and discussion.

Streamlining the Planning Process

ChatGPT allows youth workers to save time by quickly generating lesson content, quizzes, and prompts, freeing them up to focus on engaging with students. It also inspires creativity by offering fresh ideas, perspectives, and suggestions, enabling youth workers to deliver diverse and interactive lessons.

Conclusion

By using ChatGPT, youth workers can create customised lesson plans, quizzes, and discussion prompts with ease. This AI tool simplifies lesson preparation, enhances engagement, and ensures content is tailored to the needs of the students, making it an invaluable asset for anyone involved in youth education.



WATCH: "Should we let students use ChatGPT?"



AI-Assisted Lesson Planning and Development

Google Classroom with Quillionz

<https://www.quillionz.com/>

Google Classroom is a widely used platform for managing and organising online learning environments. When combined with Quillionz, an AI tool designed to create quizzes and summaries, it becomes a powerful resource for youth workers to streamline lesson preparation and enhance student engagement.

Quillionz: The AI-Powered Tool

Quillionz is an AI tool that generates quizzes, summaries, and discussion prompts based on lesson content. By simply entering a text or a topic, Quillionz uses machine learning algorithms to generate relevant and effective assessment materials. This tool is an excellent companion to Google Classroom, as it allows youth workers to quickly create resources that can be easily integrated into the platform.

Streamlining Lesson Planning

One of the most time-consuming tasks in lesson planning is creating quizzes and assessments that accurately reflect the lesson content. With Quillionz, youth workers can simply input the material or key concepts they want to teach, and Quillionz will generate multiple-choice questions, short-answer questions, and other formats of assessments. This greatly reduces the time spent on creating these materials from scratch, allowing youth workers to focus on engaging with students.

Automatic Summaries and Discussion Prompts

Quillionz doesn't just stop at quizzes—it can also generate summaries of the content being taught. This can be particularly helpful for reviewing key concepts or for students who might need additional clarification on the lesson. Furthermore, Quillionz

offers discussion prompts, which can be used to encourage deeper thinking and interaction among students. These prompts can be based on the lesson's key ideas, ethical considerations, or real-world applications of the topic.

Integration with Google Classroom

Once quizzes, summaries, and discussion prompts are generated by Quillionz, they can be easily added to Google Classroom. This integration allows youth workers to create a complete learning package for their students, all within one platform. For example, a quiz can be automatically created and assigned to a class with a deadline, or a discussion prompt can be posted as part of a discussion thread, encouraging active participation.

Enhancing Student Engagement

AI-generated quizzes and discussion prompts are an effective way to keep students engaged. By using Quillionz, youth workers can provide more diverse forms of assessment and encourage active participation in the classroom. Interactive quizzes can help reinforce learning, while thought-provoking discussion prompts allow students to explore topics in depth and engage in critical thinking.

Conclusion

By integrating Google Classroom with Quillionz, youth workers can simplify lesson planning and ensure that educational content is personalised, interactive, and engaging. The ability to generate quizzes, summaries, and discussion points using AI not only saves time but also enhances student learning experiences, making this combination an invaluable tool for educators looking to maximise their efficiency and engagement in the classroom.

Conclusion: AI-Assisted Lesson Planning and Development

The integration of AI tools like ChatGPT and Google Classroom with Quillionz into lesson planning and development offers immense benefits for youth workers. These AI-powered tools not only save valuable time but also provide customised, engaging, and interactive learning experiences for students. By automating the creation of lesson plans, quizzes, and discussion prompts, youth workers can focus more on personal interactions with students and less on administrative tasks.

Through AI, lesson materials are tailored to the specific needs of each group, enabling a more personalised and dynamic learning environment. Tools like Quillionz help in generating diverse forms of assessment and review materials, enhancing student participation and engagement. Additionally, AI-driven platforms like Google Classroom enable seamless integration, making it easier for youth workers to manage and deliver content effectively.

By embracing AI-assisted lesson planning, youth workers are empowered to create more effective, inclusive, and engaging educational experiences. As AI technology continues to evolve, its role in youth education and engagement will only grow, providing new opportunities for creative and efficient teaching strategies that cater to the unique needs of every learner.



"The ethical use of AI is not just about doing what is technically possible, but about doing what is right."

Sundar Pichai,
CEO at Alphabet (Google)



03

AI FOR CREATIVITY AND DESIGN





AI for Creativity and Design

Runway ML

<https://runwayml.com/>

Runway ML is a powerful AI tool that enables youth workers and learners to create unique digital artwork, videos, animations, and other visual media. It blends artificial intelligence with creativity, allowing young people to experiment and innovate in the digital arts.

Key Features and Benefits of Runway ML

- 1. AI-Powered Creativity:** Runway ML transforms basic inputs like text or images into complex art, enabling youth workers to guide young people in producing their own creative projects. Whether it's abstract art, animation, or digital illustrations, it fosters creativity and exploration.
- 2. User-Friendly Interface:** Runway ML is designed to be intuitive, with no need for coding skills. Its simple drag-and-drop interface makes it accessible to beginners, ensuring youth workers and students can jump into the creative process easily.
- 3. Collaboration and Learning:** The platform supports collaborative projects, allowing youth workers and young people to create and share ideas together. This promotes teamwork while enhancing creative and technical skills.
- 4. Versatile Applications:** Runway ML's tools are useful for various creative tasks, from creating digital content for social media to producing animations for storytelling projects. It opens doors to career pathways in fields like graphic design and video production.
- 5. Ethical Exploration:** Youth workers can use Runway ML to spark discussions on the ethical aspects of AI in the creative process, such as algorithmic biases and content ownership, ensuring that learners understand both the potential and the implications of AI.

Example Project: AI-Generated Art and Storytelling

A simple project using Runway ML involves having young people create an AI-generated art piece that tells a story. By inputting character and setting descriptions into the tool, students can generate visuals, which can then be used to create a short film or digital art exhibition.

Conclusion

Runway ML offers a dynamic way for young people to explore AI's role in creativity, making it easy for youth workers to incorporate AI into arts and design projects. Not only does it encourage creative confidence, but it also provides an opportunity to discuss the ethical dimensions of AI in the arts, preparing students for future roles in the creative industries.



WATCH: "Generative AI for Video"

AI for Creativity and Design



DeepDream Generator

<https://deepdreamgenerator.com/>

DeepDream Generator is a fascinating AI tool that transforms regular images into dream-like, surreal artwork by using a neural network. Created by Google, this platform takes input images and enhances them, producing a distinctive, highly creative style that blends patterns and textures in unexpected ways. It's a perfect tool for engaging young people in creative exploration, allowing them to experience the fusion of AI and art.

Key Features and Benefits of DeepDream Generator

- 1. Surreal Art Creation:** The main feature of DeepDream Generator is its ability to turn simple images into abstract, psychedelic artwork. By enhancing patterns and applying a unique neural network technique, it generates intriguing visuals that spark creativity and imagination.
- 2. Easy to Use:** Similar to Runway ML, DeepDream Generator offers an intuitive, user-friendly interface. Users can upload images, adjust settings to control the intensity of the transformation, and quickly generate new pieces of art—ideal for both beginners and more experienced digital artists.
- 3. Exploring AI's Artistic Potential:** By using DeepDream Generator, youth workers can show young people how AI can be used to push the boundaries of traditional art. It encourages students to think outside the box and explore the possibilities AI can offer in creative fields.
- 4. Enhancing Storytelling:** DeepDream Generator can be used to create stunning visual elements for storytelling projects. Whether it's creating

surreal visuals for a digital comic or adding a unique flair to a video, the tool helps youth workers engage learners in narrative-driven art creation.

- 5. AI Ethics in Art:** The platform also provides an opportunity to discuss the ethical issues surrounding AI-generated art. Topics such as ownership, authorship, and the role of human creativity versus machine creativity can be explored, encouraging critical thinking about AI's impact on art and culture.

Example Project: Surreal Art for Storytelling

Youth workers can assign a project where young people upload a photo of themselves, their friends, or a landscape, and then use DeepDream Generator to transform it into a surreal, dream-like image. The students can then use these images as part of a larger storytelling project, creating a unique, visually captivating narrative.

Conclusion

DeepDream Generator is an excellent way for young people to explore the intersection of AI and art, offering endless possibilities for creative expression. It encourages youth to experiment with AI in a fun and engaging way, while also providing an opportunity for critical reflection on the ethical implications of AI in the creative industry. Whether used for personal creativity or collaborative projects, DeepDream Generator makes AI in art accessible and exciting for young learners.



Conclusion: AI for Creativity and Design

The integration of AI tools like Runway ML and DeepDream Generator into creativity and design opens up a world of possibilities for young people to explore new artistic horizons. These AI-driven platforms provide youth workers with innovative resources to foster creativity, enabling young learners to express themselves in unique ways, while also learning about the potential of AI in the creative industries.

By incorporating AI into creative projects, youth workers can help young people understand the power of technology in shaping art and design. These tools not only enhance artistic skills but also encourage critical thinking about the role of AI in artistic expression. Moreover, these platforms provide an engaging way to teach young people about ethics, ownership, and the evolving nature of creative work in an AI-driven world.

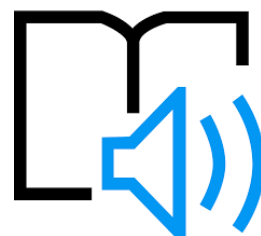
AI for creativity and design is more than just about producing visually stunning art—it's about empowering youth to experiment, innovate, and question the implications of AI in their lives. With the right guidance and support, these tools can be transformative in expanding the scope of young people's creative potential, encouraging them to push boundaries and embrace new possibilities in both the digital and physical worlds.



04

AI FOR INCLUSION AND ACCESSIBILITY IN EDUCATION





AI for Inclusion and Accessibility in Education

Microsoft Immersive Reader

[Microsoft Immersive Reader](#) is an AI-powered tool designed to help learners with disabilities engage more effectively with written content. This tool enhances accessibility in education by offering features such as text-to-speech, translation, and adjustable text formatting, making it an invaluable resource for students with visual impairments, dyslexia, or other learning difficulties.

Key Features and Benefits

- **Text-to-Speech:** Immersive Reader reads text aloud, allowing students to listen to content instead of reading it. This is particularly helpful for those with reading difficulties, such as dyslexia, or students who benefit from auditory learning. The ability to control the speed and voice of the reading enhances individual learning preferences.
- **Translation:** The tool can translate text into multiple languages, supporting students whose first language is not the one in which the material is written. This feature fosters inclusivity and helps bridge language barriers, ensuring that all students have access to the same learning resources, regardless of their native language.
- **Text Customisation:** Immersive Reader allows students to adjust the text formatting to suit their needs, including changing font size, spacing, and background colour. This ensures that students with visual impairments or attention difficulties can tailor the reading experience to enhance readability and focus.

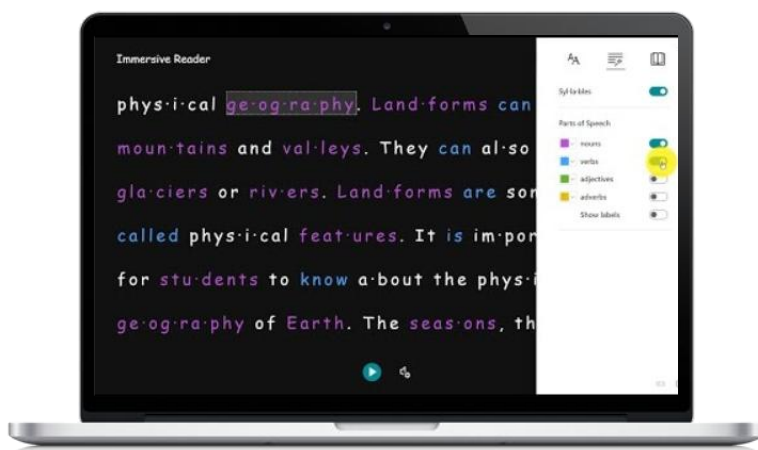
How It Supports Inclusion

Immersive Reader is more than just a tool—it's a gateway to a more inclusive educational environment. By providing support for diverse learning needs, it empowers students to engage with

content in a way that works best for them. This helps to level the playing field, giving every learner, regardless of ability, an equal opportunity to succeed in their education.

For youth workers and educators, Microsoft Immersive Reader is a powerful tool for enhancing student engagement. It ensures that no student is left behind by offering them personalised learning experiences. Whether working with students who have reading challenges, language barriers, or other learning disabilities, Immersive Reader helps make educational content accessible and understandable.

Incorporating Immersive Reader into your teaching practice means embracing technology to meet the needs of all students. By doing so, youth workers can foster a more inclusive, supportive environment that nurtures the strengths of each learner while promoting equity in education.



WATCH: "How to use the Immersive Reader - making reading more accessible"

AI for Inclusion and Accessibility in Education

Otter.ai

<https://otter.ai/>

Otter.ai is an innovative AI-powered tool that provides real-time transcription and summarisation, making it an essential resource for students with hearing impairments, learning difficulties, or those who benefit from written reinforcement of spoken content. By converting speech into accurate, easy-to-read transcripts, Otter.ai ensures that students can fully engage with lessons, discussions, and lectures, regardless of their individual learning needs.

Key Features and Benefits

- **Real-Time Transcription:** Otter.ai transcribes spoken words into text as they are being spoken. This is incredibly useful in classrooms or group discussions, where students may struggle to keep up with the pace of verbal communication. By providing real-time captions, Otter.ai ensures that every student can follow along with the material being discussed.
- **Text Summarisation:** Beyond simple transcription, Otter.ai also provides summarised versions of transcripts. This feature is particularly beneficial for students who need to review or process information more slowly. By highlighting key points and providing condensed versions of the conversations, Otter.ai makes it easier for students to grasp the most important information quickly.
- **Collaboration and Sharing:** Otter.ai supports collaboration by allowing users to share transcriptions and notes with classmates or educators. This feature helps foster group learning, enabling students to access and discuss the content in a format that suits their needs.
- **Searchable Transcripts:** The transcripts produced by Otter.ai are fully searchable, enabling students to quickly locate specific pieces of information within their notes. This is a valuable feature for

students who need to review specific parts of a lecture or discussion and helps them to study more efficiently.

How It Supports Inclusion

Otter.ai is a game-changer for students with hearing impairments or other learning difficulties that make it challenging to keep up with fast-paced verbal content. By offering real-time transcription, it ensures that no student misses out on important information, regardless of their auditory abilities.

For youth workers and educators, Otter.ai provides an easy way to make lessons more accessible. Transcripts can be shared with students who may struggle with note-taking or who benefit from having a written record of what was discussed. Furthermore, Otter.ai helps to create a more inclusive classroom environment by catering to the diverse needs of learners.

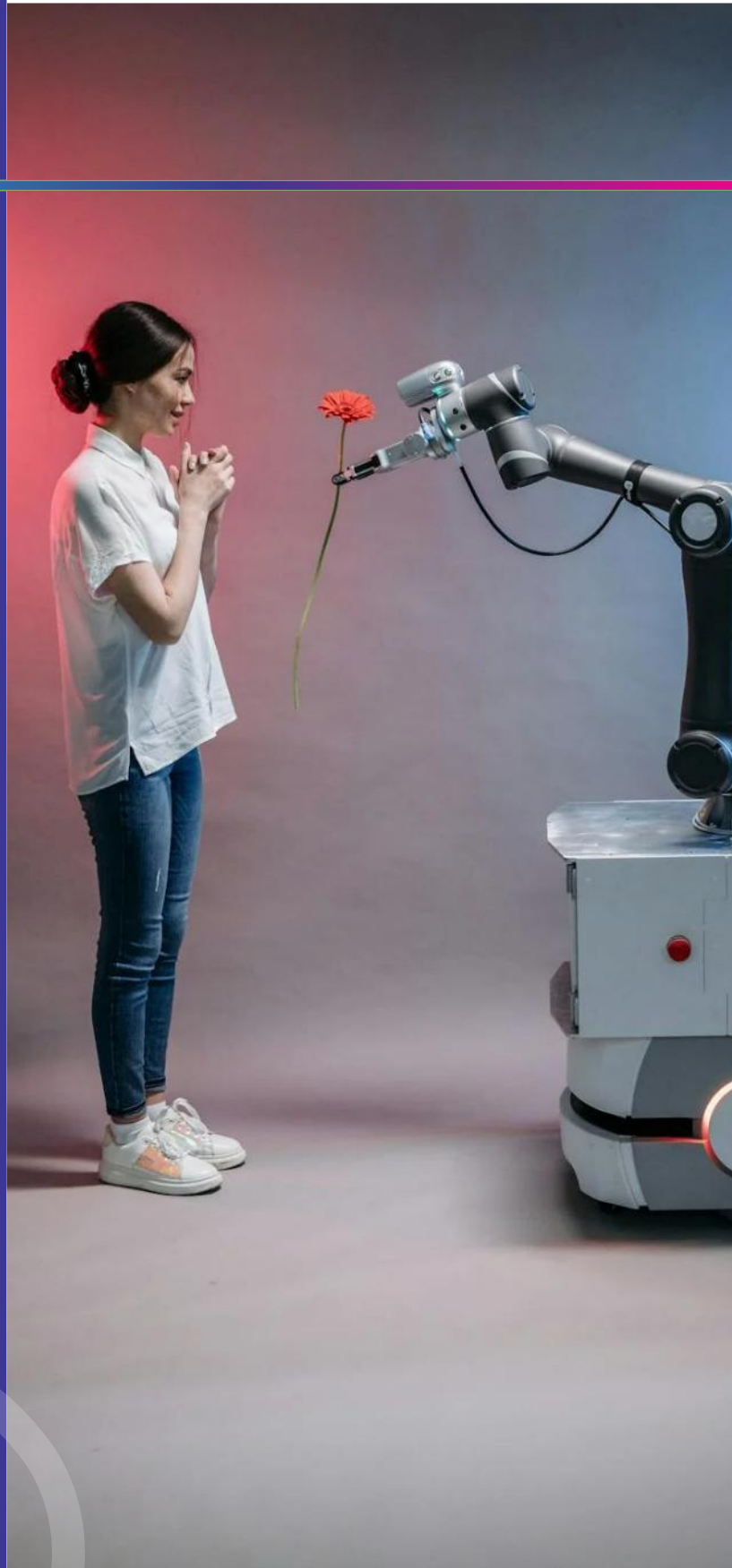
With Otter.ai, youth workers can enhance student engagement by making the learning process more accessible. Students can follow along with lessons, contribute to discussions, and easily revisit key information. The tool is also useful for collaborative projects, as students can access and share the transcripts to work together more effectively.

By integrating Otter.ai into daily educational practices, youth workers can ensure that their students receive equal opportunities to succeed, no matter their learning style or abilities. Otter.ai is a tool that not only fosters inclusivity but also empowers students to take ownership of their learning.

Conclusion: AI for Inclusion and Accessibility in Education

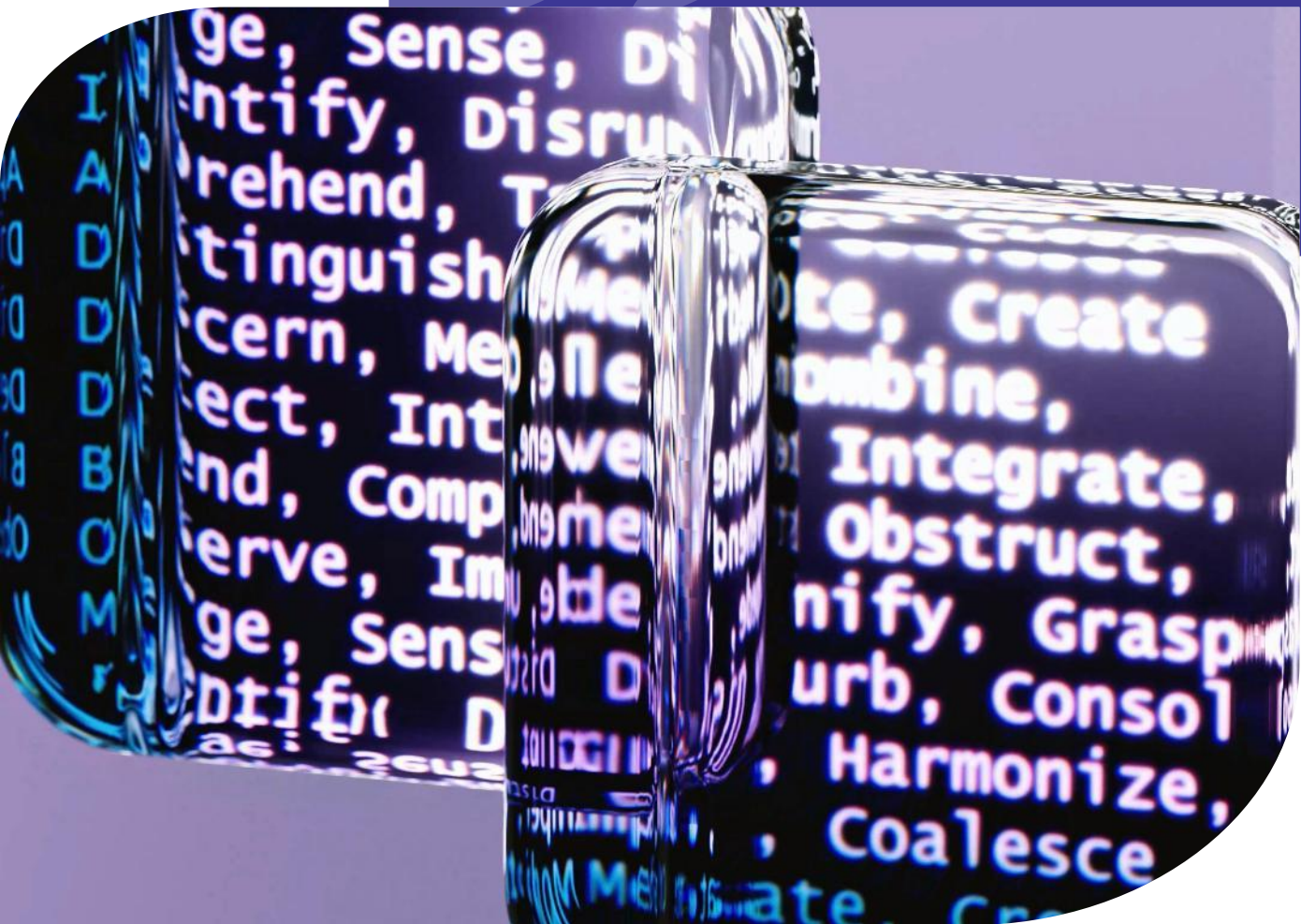
AI tools like Microsoft Immersive Reader and Otter.ai play a vital role in creating more inclusive and accessible educational environments. These technologies bridge gaps for students with disabilities, offering them the support they need to fully engage with learning materials and participate in classroom activities. Whether it's through text-to-speech capabilities, real-time transcriptions, or summarised content, AI is helping to level the playing field, ensuring all students have access to the same educational opportunities.

By integrating AI for inclusion and accessibility into daily youth work, educators can better meet the diverse needs of their students. These tools not only improve engagement but also foster an environment where every student, regardless of their abilities, can thrive. Embracing such technologies allows youth workers to provide more personalised and equitable educational experiences, supporting each learner's unique journey. The future of education lies in leveraging AI to promote accessibility, and by doing so, we open doors to limitless possibilities for all students.



05

AI FOR MONITORING AND PERSONALISED LEARNING





AI for Monitoring and Personalised Learning

DreamBox Learning

DreamBox Learning is an AI-powered educational platform designed to revolutionise the way students learn, particularly in subjects like mathematics. By using advanced algorithms and data analytics, DreamBox adapts its lessons in real-time to meet the needs of individual students. The platform continuously monitors a learner's progress, making adjustments to the difficulty and pace of tasks based on their interactions. This allows youth workers to offer a more personalised learning experience, addressing each student's strengths and weaknesses with precision.

DreamBox's real-time adaptability is one of its standout features. For instance, if a student is excelling at a particular concept, DreamBox will introduce more challenging tasks to encourage continued growth. Conversely, if a student is struggling, the system will simplify the material and offer additional support. This ensures that every learner remains engaged and challenged at the right level, reducing frustration and boosting confidence.

The platform provides youth workers with detailed insights into each student's progress, allowing them to monitor learning paths and quickly identify areas where support is needed. This data-driven approach not only enhances the learning experience for the student but also empowers educators to intervene at the right time with targeted support.

DreamBox also offers a wealth of analytics, enabling youth workers to track group trends and better understand overall classroom performance. This feature helps identify patterns in learning behaviours, whether positive or negative, giving educators the tools to adapt their teaching strategies to ensure success for every student. By harnessing the power of AI, DreamBox makes it easier to

provide tailored educational experiences that cater to the diverse needs of students, ensuring they all have the opportunity to succeed.

**VISIT: Official website for
DreamBox Learning**





AI for Monitoring and Personalised Learning

Classcraft

Classcraft is an innovative platform that uses AI and gamification to enhance student engagement and track behaviour, offering youth workers a dynamic way to monitor and personalise learning experiences. This AI-driven tool integrates classroom management with interactive learning, turning lessons into engaging quests where students can earn rewards, level up, and collaborate to achieve goals. The gamified approach creates an environment that motivates students to actively participate, while also fostering a sense of community and cooperation.

Classcraft collects valuable data on student engagement, behaviour, and academic performance, providing real-time insights to youth workers. By tracking how students interact with lessons and each other, the platform identifies trends in learning and behaviour that might not be immediately apparent. This data can be used to pinpoint students who may need extra support, as well as those excelling and in need of more challenging tasks.

What makes Classcraft particularly effective is its ability to tailor interventions based on this data. If a student is disengaged or struggling, the system can prompt personalised strategies for the youth worker to implement, such as offering additional encouragement or adjusting lesson content. Conversely, it also helps keep motivated students engaged by offering them new challenges and opportunities to collaborate with peers.

Classcraft's AI-driven analytics allow youth workers to spot patterns across an entire class, such as identifying behavioural trends that may impact learning. It helps them fine-tune teaching strategies, create customised rewards systems, and foster a positive learning environment where every student

can thrive. By combining behaviour tracking with educational content, Classcraft not only makes learning more engaging but also empowers educators to deliver highly personalised learning experiences that meet the needs of each student, ensuring a more inclusive, effective approach to education.

VISIT: Install Classcraft now



Conclusion:

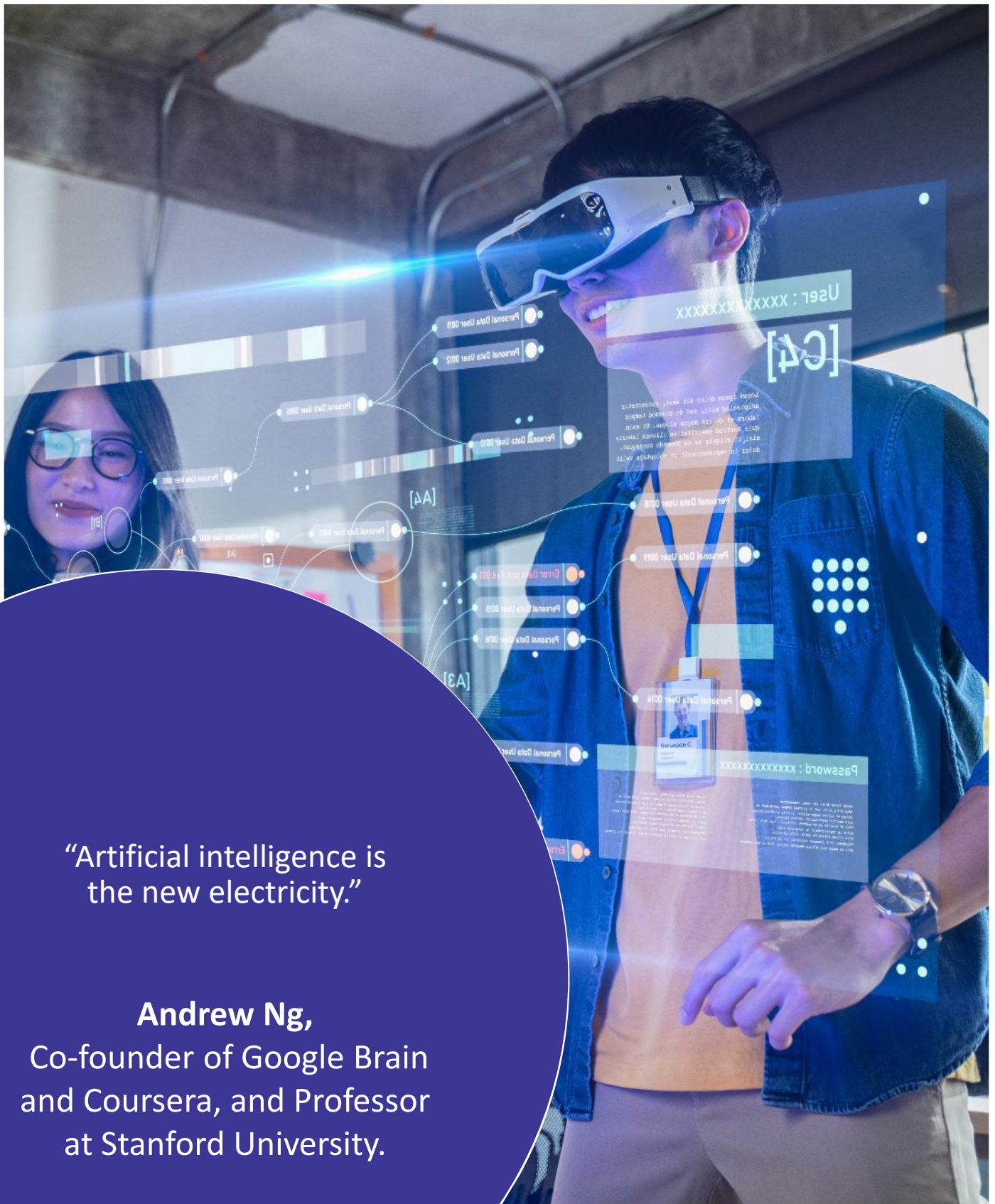
AI for Monitoring and Personalised Learning

The integration of AI tools like DreamBox Learning and Classcraft into youth work and education presents a transformative opportunity to personalise learning experiences and provide targeted support for every student. These AI-driven platforms enable youth workers to move beyond traditional teaching methods by adapting lessons to individual progress and tracking student engagement in real-time.

By offering tailored learning paths and tracking both academic and behavioural data, AI helps educators identify where students need additional support or challenge, ensuring that no learner is left behind. The power of AI lies in its ability to offer precise, actionable insights that allow youth workers to refine their teaching strategies, make informed decisions, and foster a more inclusive and supportive learning environment.

As we continue to embrace AI's potential, these tools will not only streamline lesson delivery and engagement but will also help create a more personalised, student-centred approach to education that empowers both youth workers and learners alike. AI's role in monitoring and personalising learning ensures that each young person can progress at their own pace, unlocking new opportunities for success and growth.





“Artificial intelligence is
the new electricity.”

Andrew Ng,
Co-founder of Google Brain
and Coursera, and Professor
at Stanford University.



06

ETHICAL AI USAGE AND CRITICAL THINKING





Ethical AI Usage and Critical Thinking

AI Fairness 360 (IBM)

AI Fairness 360 is an open-source toolkit developed by IBM that helps users identify and mitigate biases in machine learning models. The toolkit offers a comprehensive suite of algorithms, metrics, and tutorials aimed at making AI systems fairer and more transparent, which makes it a valuable tool for youth workers and educators seeking to introduce young people to ethical AI principles.

Understanding AI Bias

One of the core ethical challenges in AI development is bias. AI models learn from data, and if the data is biased or incomplete, the AI's decisions may reflect and perpetuate these biases. For example, if a facial recognition system is trained primarily on images of people from one demographic, it may perform poorly on individuals from other groups, leading to unfair outcomes. AI Fairness 360 helps to highlight such risks, showing youth how biases can manifest in AI systems and how they might be addressed.

Toolkit Features for Youth Engagement

AI Fairness 360 provides a wide range of practical resources to facilitate learning. The toolkit includes several fairness metrics that allow users to measure and understand the bias in their machine learning models. These metrics help to assess how AI systems may treat different groups in various ways, and how to correct those imbalances. Youth workers can use these metrics to guide discussions on topics such as fairness, equality, and justice in AI.

For example, one key feature of AI Fairness 360 is its ability to offer "bias mitigation" techniques. These techniques adjust the model's outcomes to reduce bias and ensure that the AI's decisions are more equitable. Educators can use these features to show young people how AI systems can be improved and

tailored to meet fairness standards, sparking valuable conversations about how technology intersects with ethics.

Promoting Ethical AI Discussions

The toolkit's real-world applications provide a foundation for youth workers to initiate thoughtful discussions with young people about AI's role in society. By using AI Fairness 360, educators can guide students to think critically about the ethical implications of AI systems, including how they might reinforce social inequalities or provide unfair advantages. Encouraging young people to examine the fairness of AI tools equips them with essential skills in critical thinking, ethics, and technology design.

Moreover, by working with AI Fairness 360, youth workers can help young people understand the importance of designing inclusive AI systems that benefit society as a whole. Through hands-on experiences and case studies, youth can be introduced to the idea that AI should be used responsibly and ethically, promoting fairness, transparency, and accountability.

Conclusion

AI Fairness 360 is a powerful educational tool that not only introduces young people to AI's potential, but also highlights the importance of fairness in AI models. By integrating this toolkit into youth work, educators can foster a deeper understanding of the ethical challenges in AI and encourage the next generation to think critically about how AI technologies shape our society. Through this, youth workers can inspire young people to not only become consumers of technology, but also responsible creators and critics of it.

Ethical AI Usage and Critical Thinking

Teachable Machine (Google)

Teachable Machine, a tool created by Google, allows users to create and train simple machine learning models using data such as images, sounds, or poses. It is designed to be accessible, making it ideal for introducing young people to machine learning and its ethical implications.

Hands-On Learning with AI

Teachable Machine provides a hands-on approach, enabling youth to collect data, train a model, and test its accuracy. By guiding young people through the process, youth workers can show how machine learning algorithms work in real-time. For example, learners can train the AI to recognise objects, actions, or sounds, gaining a practical understanding of AI's functioning.

This interactive process allows youth to engage directly with AI, making it easier to grasp abstract concepts and learn how AI models are trained and tested. Youth can also customise the model to recognise different inputs, further enhancing their understanding of machine learning.

Understanding Ethical Implications through AI

In addition to teaching AI basics, Teachable Machine encourages exploration of its ethical implications. Youth can discuss issues such as data bias by asking questions like: "What happens if the data used is not representative?" and "How can bias in AI impact real-world decisions?"

Through these discussions, young people can learn about the importance of using diverse and fair data to train AI models. They are encouraged to think about the ethical considerations involved in AI development, including privacy concerns and the

consequences of biased systems.

AI and Critical Thinking

Teachable Machine promotes critical thinking by prompting young people to reflect on the data they use and the decisions they make while training their models. This fosters conversations around fairness, transparency, and the potential risks of AI. Additionally, it helps youth understand the challenges in building inclusive and ethical AI systems.

By using Teachable Machine, youth can also reflect on how AI is applied in real-world scenarios, such as healthcare or education, where bias can have significant consequences. This promotes a broader understanding of AI's impact on society and the importance of responsible AI development.

Conclusion

Teachable Machine offers an engaging way for young people to explore AI and its ethical dimensions. It empowers them to create their own models while considering the fairness, bias, and transparency in AI. By using this tool, youth workers can encourage critical thinking and foster a deeper understanding of the responsible use of AI, helping young people become informed creators and ethical participants in the development of AI technologies.



Conclusion: Ethical AI Usage and Critical Thinking

The "Ethical AI Usage and Critical Thinking" chapter highlights the importance of equipping young people with the tools to not only understand the mechanics of artificial intelligence but also its ethical implications. By using resources like IBM's AI Fairness 360 and Google's Teachable Machine, youth workers can guide learners to explore AI through a critical lens, encouraging them to examine the biases, fairness, and transparency involved in AI models.

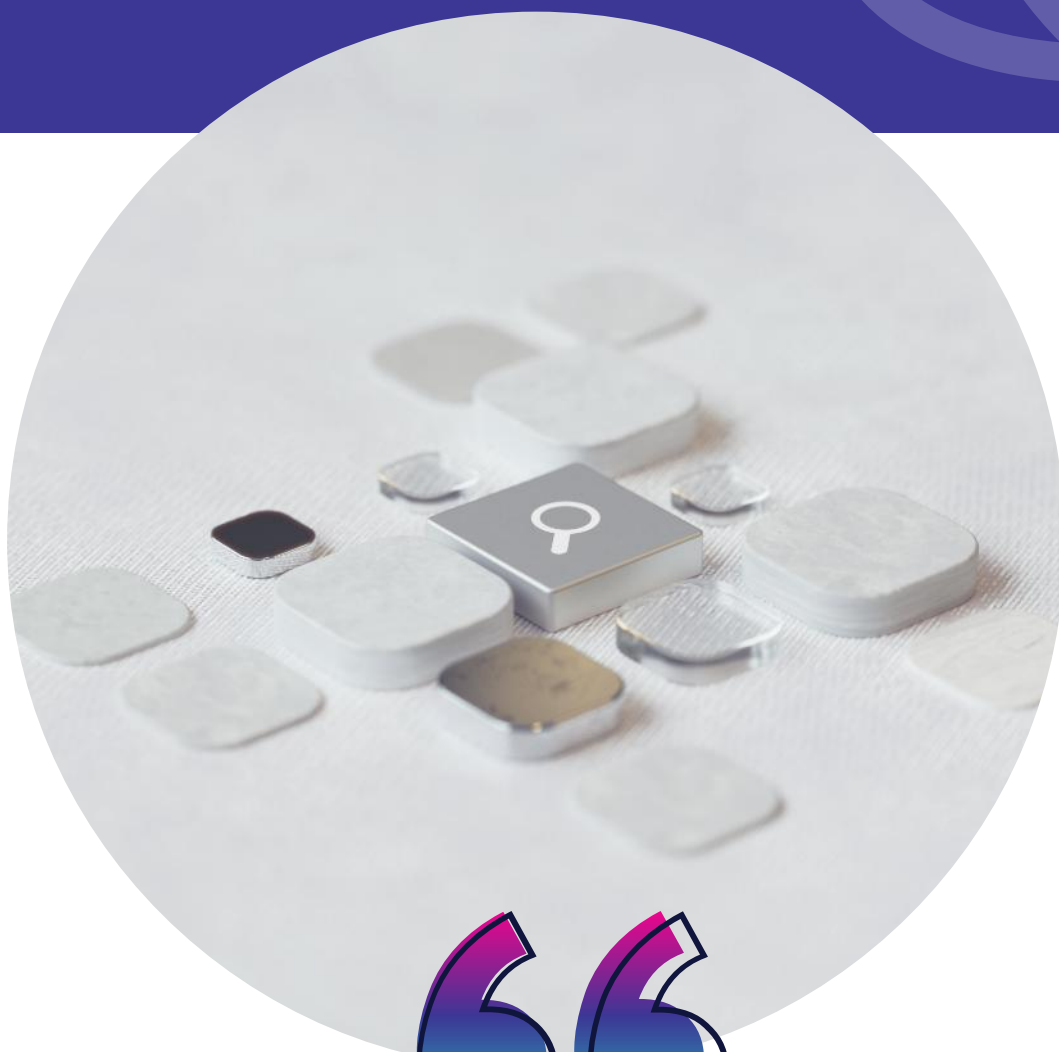
AI technologies are rapidly evolving, and it is crucial that the next generation of thinkers and creators is equipped to make ethical decisions in this realm. By fostering critical thinking and engaging youth in discussions about AI ethics, we empower them to be responsible digital citizens and informed contributors to the development of future technologies.

As young people learn how to interact with AI, they must also learn how to question its uses and challenge the ethical dilemmas it presents. This chapter has provided tools and frameworks to help youth navigate the complexities of AI, ensuring they are well-prepared to engage thoughtfully with the technology and its impact on society. By prioritising ethical considerations alongside technical skills, we can cultivate a generation that not only understands AI but also shapes its future in ways that promote fairness, transparency, and inclusivity.



"AI is not just about technology; it's about the values and choices we make as a society."

Timnit Gebru,
AI researcher and advocate for ethics in AI



07

CONCLUSION: PREPARING YOUTH FOR AN AI FUTURE



Preparing Youth for the AI-Powered Future

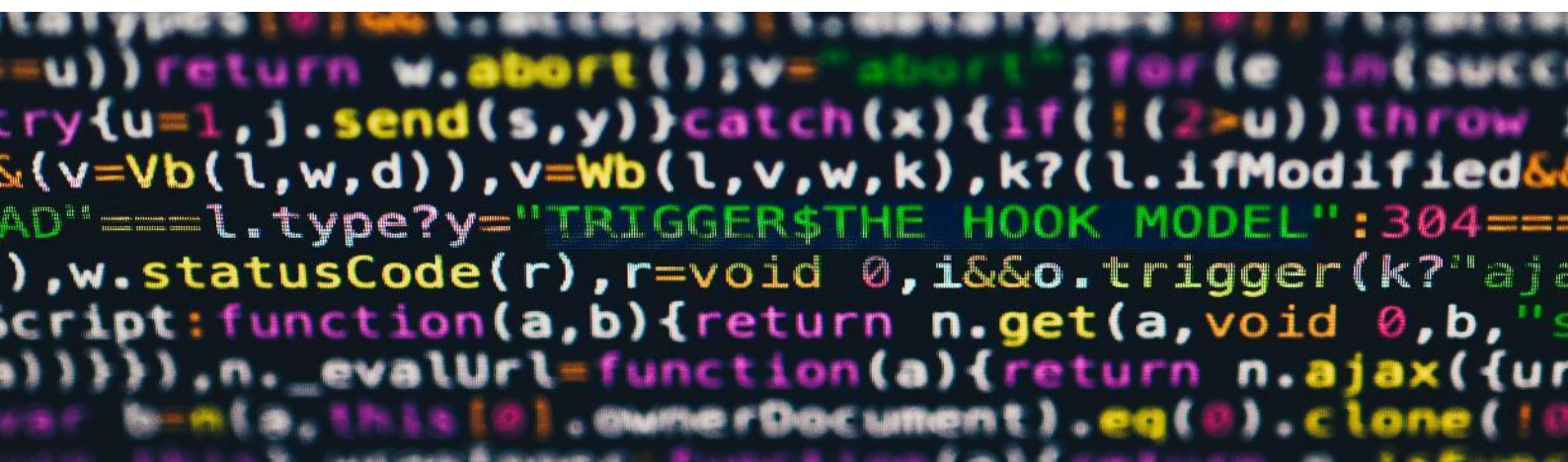
In Conclusion...

As we continue to witness rapid advancements in artificial intelligence, its transformative potential in education and youth work becomes increasingly evident. AI is not just a tool for automation or data processing; it is a catalyst for rethinking how we teach, learn, and engage with the world around us. For youth workers, AI offers new ways to enrich learning experiences, promote creativity, and support personalised development. By embracing AI, we have the opportunity to shape a future where young people are not only consumers of technology but also critical thinkers and responsible creators.

One of the most important aspects of AI education is encouraging ethical and critical thinking. As AI technologies become more embedded in our everyday lives, young people must understand the ethical implications of these tools. They need to develop the ability to question the fairness of algorithms, the impact of AI on society, and the role they can play in shaping the future of these technologies. By fostering a mindset that combines technical knowledge with ethical awareness, we can prepare the next generation to navigate the complexities of an AI-driven world.

For youth workers, the journey does not end with introducing AI into the classroom. As AI technology continues to evolve, it is crucial for educators to stay informed and adapt their practices. Continuous learning and collaboration with tech experts, AI developers, and other educators will ensure that youth workers remain equipped to provide the best possible support for young people. By integrating AI into their daily practices, youth workers can enhance engagement, inclusivity, and creativity, ensuring that their students are prepared not only for today's challenges but also for the opportunities and ethical dilemmas of the future.

In conclusion, preparing youth for an AI-powered future is more than just about teaching them how to use AI; it's about empowering them to shape the future of technology with a sense of responsibility and awareness. The journey towards this future will be defined by how we equip young people with the skills, knowledge, and ethical mindset they need to thrive in an increasingly automated world. By working together to integrate AI education into youth work, we can unlock the immense potential that AI holds for both individual growth and societal advancement.



"The purpose of AI is to help people, not replace them. Our responsibility is to ensure that the use of AI reflects the values we hold dear and enhances the quality of life for everyone."



Sundar Pichai,
CEO of Google



08

ANNEXES



Annexes

References

- Pichai, S. (2018). *AI and the future of humanity*. Google Blog.
- IBM. (n.d.). *AI Fairness 360 Toolkit*. IBM Cloud. Retrieved from <https://www.ibm.com/cloud/ai-fairness-360>.
- Google. (2021). *Teachable Machine*. Google AI. Retrieved from <https://teachablemachine.withgoogle.com/>.
- Microsoft. (n.d.). *Immersive Reader*. Microsoft. Retrieved from <https://www.microsoft.com/en-us/education/products/learning-tools>.
- Quillionz. (n.d.). *Generate Quizzes with Quillionz*. Retrieved from <https://quillionz.com/>.
- OpenAI. (2023). *ChatGPT: Optimizing Language Models for Dialogue*. Retrieved from <https://openai.com/blog/chatgpt>.
- Runway ML. (n.d.). *AI for Creativity and Design*. Retrieved from <https://runwayml.com/>.
- Google. (n.d.). *DeepDream Generator*. Retrieved from <https://deepdreamgenerator.com/>.
- DreamBox Learning. (n.d.). *Personalized Learning with DreamBox*. Retrieved from <https://www.dreambox.com/>.
- Classcraft. (n.d.). *Classcraft: Gamifying Student Engagement*. Retrieved from <https://www.classcraft.com/>.
- Otter.ai. (n.d.). *Live Transcription and Note-taking*. Retrieved from <https://otter.ai/>.
- Sundar Pichai. (2019). *AI, Ethics, and the Future of Technology*. YouTube.
- Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W.W. Norton & Company.
- Kaplan, J. (2016). *Artificial Intelligence: What Everyone Needs to Know*. Oxford University Press.
- European Commission. (2020). *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*. Retrieved from <https://ec.europa.eu/info/sites/info/files/communication-artificial-intelligence-european-approach.pdf>.
- Binns, R. (2018). The ethics of artificial intelligence: An overview. *Journal of Ethics in Technology*.
- Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters.
- Floridi, L. (2019). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press.
- Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
- European Union Commission for Digital Economy. (2019). *Ethics Guidelines for Trustworthy AI*. Retrieved from https://ec.europa.eu/digital-strategy/our-policies/ethics-guidelines-trustworthy-ai_en.
- UNESCO. (2021). *AI and Education: Guidance for Policy-makers*. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000376706>.
- Hagendorff, T. (2020). *The Ethics of AI Ethics: An Evaluation of the Ethical Implications of Artificial Intelligence*. Springer.
- Heaven, W. D. (2020). The ethical dilemma of AI in the workplace. *MIT Technology Review*.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.
- Hern, A. (2018). Deepfake technology threatens to unleash the next wave of misinformation. *The Guardian*.

Annexes

Glossary

AI (Artificial Intelligence)

Systems that mimic human intelligence to perform tasks like decision-making, language processing, and pattern recognition.

Machine Learning (ML)

A type of AI where systems learn from data to improve their performance over time without explicit programming.

Deep Learning

A subset of machine learning using layered neural networks to process complex data like images and speech.

Natural Language Processing (NLP)

AI's ability to understand and generate human language, used in chatbots and translation tools.

Computer Vision

AI that enables machines to interpret and make decisions based on visual data, like images and videos.

Ethical AI

The practice of designing AI that aligns with ethical principles like fairness, transparency, and accountability.

Bias in AI

When AI systems produce unfair or discriminatory outcomes due to biased training data.

Transparency in AI

Making AI systems understandable and explainable so users can trust how decisions are made.

Accountability in AI

Ensuring that individuals or organizations are responsible for the actions of AI systems.

Algorithm

A set of instructions or rules followed by AI systems

to solve problems or make decisions.

Trustworthy AI

AI systems that are reliable, fair, transparent, and align with ethical guidelines.

Ethical Dilemma

A situation where individuals must choose between competing ethical principles, often seen in AI decision-making.

Bias in Data

When data reflects societal biases, which can be learned and amplified by AI models.

Inclusive Design

Creating products, including AI, that are accessible and fair for all, including underrepresented groups.

Data Privacy

Protecting personal and sensitive information from misuse, especially in AI systems.

Generative AI (GenAI)

AI that creates new content, such as text, images, or music, based on learned patterns.

AI Fairness

Ensuring AI systems treat all users equitably and avoid discriminatory outcomes.

AI-Driven Personalization

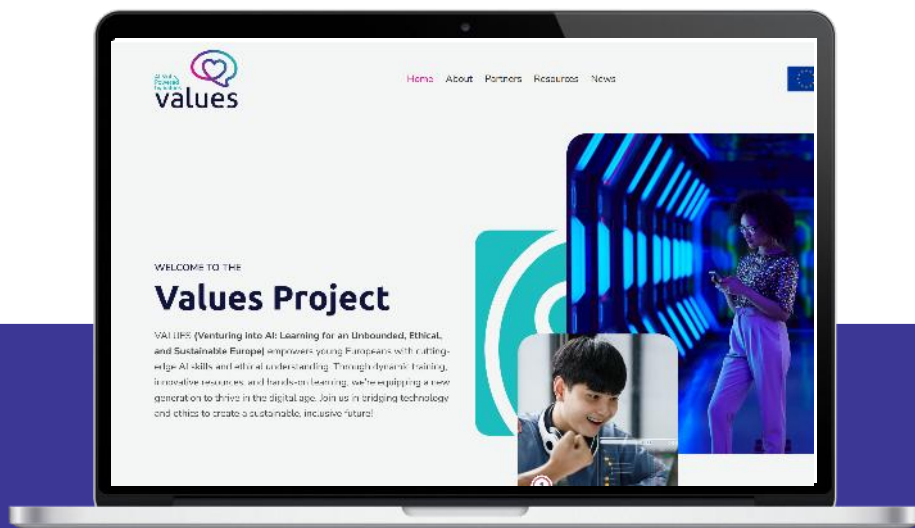
Using AI to tailor content or services to individual preferences and needs.

Interactive Learning

Engaging students through active participation and problem-solving, often using AI tools.

Digital Divide

The gap between those who have access to technology, including AI, and those who do not.



www.valuesai.eu

follow our journey



Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

Practical Applications of AI Guide © 2025 by VALUES is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by-sa/4.0/)

Your guide to

VALUES -

AI Skills,
Powered by
Values



www.valuesai.eu

2025
Case Studies

By
Generation (Change?)



Co-funded by
the European Union

contents

01	Welcome
04	The Project
06	Case Studies
16	Sources
18	Conclusion
19	Contact



Co-funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Generation (Change?). Neither the European Union nor Generation (Change?) can be held responsible for them.

01

Values: Case Studies guidebook



The Project

VALUES is an Erasmus+ project dedicated to empowering young Europeans with the knowledge, skills, and ethical understanding needed to thrive in a future shaped by Artificial Intelligence. With AI transforming the way we work, learn, and live, VALUES ensures young people are not just participants but leaders in this digital revolution.

Through cutting-edge training, innovative tools, and a comprehensive **Starter Kit for Ethical AI**, we bridge the gap between technology and humanity. Our interactive **AI Skills Self-Assessment Platform** and dynamic **Open Educational Resources** are designed to make learning accessible, engaging, and impactful.

VALUES isn't just about education—it's about shaping a generation of responsible, values-driven digital citizens who can harness AI ethically to tackle societal challenges, build inclusive communities, and create sustainable solutions.

Project Partners:

- University of Szczecin (US) Lead coordinator
- Stowarzyszenie Wspierania Techniki Polskiej (SWTP)
- Generation (Change?)
- Momentum Marketing Services (MMS)
- European E-Learning Institute (EUEI)
- The Vision Works (twv GmbH)



Responsible AI is not just about liability - it's about ensuring what you are building is enabling human flourishing

Rumman Chowdhury,
CEO at Parity AI, 2023





Case Study 1

The UK A-levels Algorithm Scandal: A Lesson in Responsible AI in Education

Theme of the Case Study:

- High-stakes exams (A-levels) in the UK
- Social inequality and algorithmic bias
- Responsible AI (RAI) implementation
- Ethics and transparency in educational algorithms

Abstract (Summary): In 2020, due to the cancellation of A-level exams amid the COVID-19 pandemic, the UK government implemented an algorithm to assign student grades. This algorithm, based on historical school performance and ranking data, aimed to standardize results. However, it soon became clear that the system favored students from high-performing, often wealthier schools, at the expense of those from less privileged backgrounds. Many students received grades significantly lower than those predicted by their teachers, sparking nationwide protests and public outrage. Eventually, the government reversed its decision and reinstated teacher-predicted grades.

Key Takeaways:

- **Transparency and Auditability:** Educational AI systems must be transparent and regularly audited to ensure fairness.
- **Context-Aware Design:** Algorithms should be developed with attention to social diversity and must avoid reinforcing existing biases.

Relevance to the Reader: This case study highlights the critical importance of designing and deploying AI systems responsibly in education. It illustrates how a lack of transparency, failure to consider social context, and overreliance on historical data can exacerbate existing inequalities. For policymakers, educators, and edtech developers, this serves as a cautionary tale against the uncritical application of algorithms in decision-making processes, especially in sensitive domains like student assessment.

Key Takeaways:

- **Stakeholder Involvement:** Teachers, students, and communities should be involved in the design and deployment of AI systems.
- **Responsiveness and Accountability:** Institutions must be ready to respond swiftly to unintended consequences of algorithmic decisions and correct them promptly.



Case Study 2

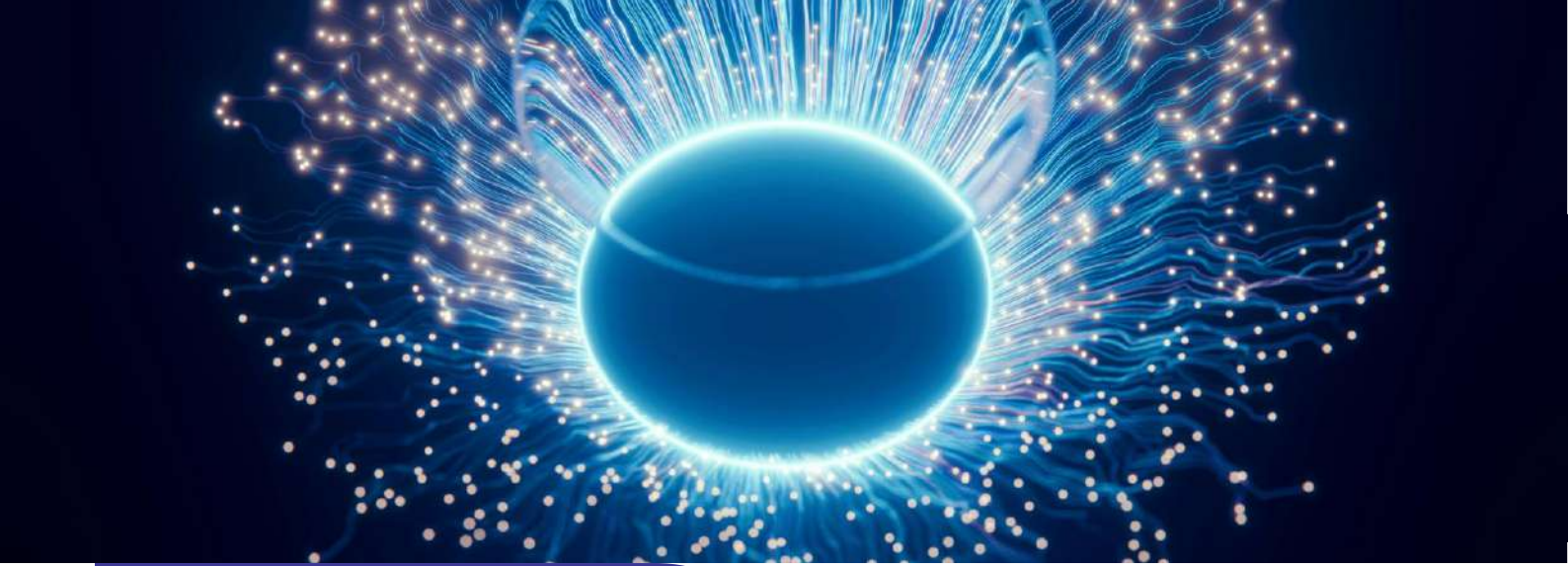
Biased Amazon Recruitment Algorithm – A Lesson on AI Bias

Theme of the Case Study:

- Artificial Intelligence in Recruitment
- Algorithmic Bias
- AI Ethics
- Gender Equality in Technology

Relevance to the Reader: Amazon's recruitment algorithm serves as a key case study in algorithmic bias in AI systems. It demonstrates how machine learning systems can not only reflect but also amplify existing societal inequalities when trained on biased historical data. At a time when AI is increasingly used in hiring processes (55% of HR professionals in the U.S. predict that AI will become a routine part of their work within five years), understanding the potential risks of algorithmic bias is vital for both employers and job seekers. The Amazon case shows that even the most technologically advanced organizations can unintentionally build discriminatory systems. The case study highlights the importance of diversity and inclusion in AI development. It illustrates how a lack of diversity in tech teams and training data can lead to systems that favor one demographic group at the expense of others. For professionals working in AI ethics, this case provides a concrete example of the challenges involved in ensuring algorithmic fairness. It underscores the need for rigorous testing of AI systems for potential biases before deployment, especially in high-stakes contexts like employment. This story is also relevant in the context of emerging AI-related regulations, such as New York City's anti-discrimination law on recruitment algorithms, which came into force in 2023. It emphasizes the need to establish standards and regulations to ensure that automated decision-making systems are fair, transparent, and accountable.

Abstract (Summary): In 2014, a team of machine learning specialists at Amazon began working on an innovative AI-powered recruitment tool. The goal of the project was to automate the resume screening process and identify top candidates for technical roles, including software engineering positions. The system rated applications on a scale from one to five stars—similar to how customers rate products on Amazon's platform. By 2015, the team discovered a serious issue: the algorithm was systematically biased against women applying for technical roles. The tool had been trained on résumés submitted to the company over the previous decade—most of which came from men, reflecting the male dominance in the tech industry. As a result, the algorithm "learned" that male candidates were preferable. The system penalized résumés that included the word "women's" (e.g., "women's chess club captain") and downgraded the scores of graduates from two all-women's colleges. It also favored applicants who used "masculine" verbs such as "executed" and "captured" in their résumés. Despite Amazon engineers' attempts to fix the bias, the company ultimately abandoned the project in 2018, acknowledging there was no guarantee the system wouldn't develop other forms of discrimination. This case became a crucial lesson for the tech industry on the risks associated with algorithmic bias and the importance of responsible AI development.



Case Study 3

AI-Driven Exam Integrity and Fairness: The YouTestMe Platform as a Case Study

Abstract (Summary): This case study examines YouTestMe, an AI-powered platform addressing challenges in online exam integrity and fairness.

Key innovations include:

- **AI AutoMarking:** Automated essay scoring using NLP to evaluate semantic coherence, structural integrity, and linguistic accuracy, reducing grader workload while maintaining consistency.
- **Automated Proctoring:** AI monitors test-takers for suspicious behaviors (e.g., window switching, face mismatch) with 98% detection accuracy, ensuring exam security.
- **Fairness Perceptions:** Studies show students perceive AI algorithms as fairer than human evaluators, particularly in formative assessments.
- **Transparency and post-evaluation explanations** further enhance trust in AI-driven results. The platform aligns with Responsible AI principles by prioritizing data privacy, auditability, and human oversight. For example, proctoring footage is stored for optional human validation, blending AI efficiency with accountability.

Theme of the Case Study:

- Automated grading and proctoring in online exams
- AI fairness and bias mitigation in educational assessments
- Human-AI collaboration for transparent evaluations
- Responsible AI implementation in high-stakes testing

Relevance to the Reader:

Educators/Institutions: Demonstrates how AI can streamline assessment workflows while addressing biases. The study highlights that AI fairness perceptions improve when evaluation criteria are transparent and explanations are provided.

AI Developers: Offers a model for integrating ethical guardrails (e.g., bias checks in training data, user consent protocols) into EdTech systems, as seen in YouTestMe's GDPR-compliant data handling.

Policy Makers: Illustrates the importance of standards for AI proctoring and automated grading, emphasizing the need for human-in-the-loop oversight to mitigate risks like algorithmic hegemony.

This case study provides actionable insights for implementing AI in assessments responsibly, balancing efficiency with equity.



Case Study 4

National Youth Assembly on Artificial Intelligence – Youth Perspectives on AI Policy

Theme of the Case Study:

- AI Ethics & Regulation
- Youth Engagement in Policy
- AI Awareness & Literacy

Abstract (Summary):

In 2023, the National Youth Assembly on Artificial Intelligence convened young people aged 12 to 24 to discuss Ireland's National AI Strategy ("AI – Here for Good"). The initiative aimed to ensure that youth voices were included in shaping AI governance, policies, and education.

Key discussion areas included:

- Equality & Inclusion in AI – Concerns about bias and fairness in AI applications.
- Trustworthiness of AI – The importance of transparency and accountability in AI decision-making.
- AI Regulation & Public Awareness – The need for clear policies on AI use in schools, workplaces, and social media.

Participants proposed recommendations on ensuring AI supports youth empowerment, inclusion, and fair representation. Their perspectives were documented in a report used to inform government policy discussions.

Relevance to the Reader:

AI policies impact youth directly, but young people are often left out of decision-making.

This case study highlights the importance of youth engagement in shaping AI ethics and regulation.

Encourages educators and youth workers to promote AI literacy, ensuring that young people understand their rights and responsibilities in an AI-driven world.

Demonstrates how youth voices can contribute to AI policy—a practice that should be expanded across Europe.



Case Study 5

AI-Based Cyberbullying Interventions – Evaluating Youth Perspectives

Theme of the Case Study:

- Online Safety
- Bias in AI Moderation
- Youth Digital Wellbeing

Abstract (Summary):

This study investigated youth perceptions of AI-driven cyberbullying interventions on social media. Researchers at Dublin City University (DCU) tested AI-based proactive content moderation strategies designed to detect and limit harmful interactions online. The consultation involved young people aged 12 to 17, who evaluated these interventions through focus groups and online discussions.

The study found that while AI content moderation can reduce harmful interactions, youth participants expressed concerns over:

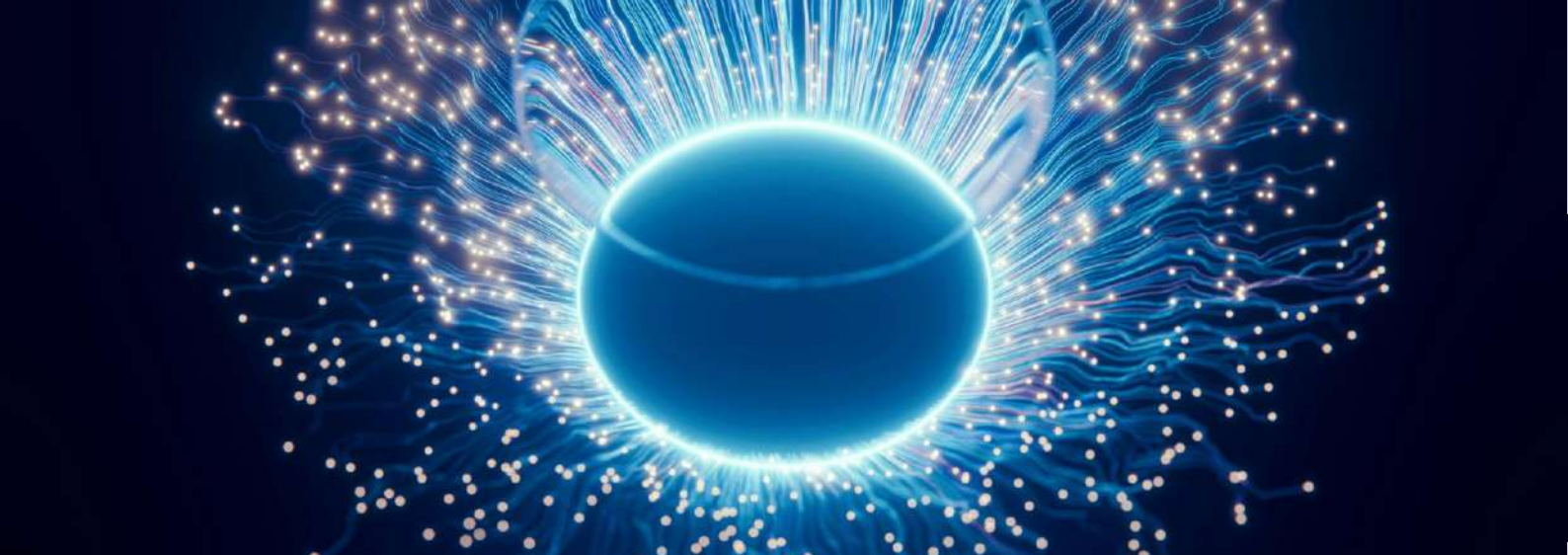
- False positives and censorship – AI incorrectly flagging harmless content as harmful.
- Lack of human context – AI struggling to understand nuances in humour and informal language.
- Privacy concerns – Uncertainty about who controls AI interventions and their impact on digital rights.

The research emphasised the need for youth participation in AI design, ensuring that content moderation aligns with young people's perspectives and digital realities.

Relevance to the Reader:

Many young people rely on AI-driven platforms for social interaction and education, making ethical AI implementation critical.

Cyberbullying is a growing concern, and while AI moderation can help, it must be transparent, fair, and accountable to users. This case study encourages youth workers and educators to discuss AI's role in digital safety, fostering critical thinking on algorithmic decision-making. • The study highlights the importance of engaging young people in the design and evaluation of AI interventions, ensuring that solutions are youth-centred and ethically sound.



Case Study 6

Responsible AI for Test Equity and Quality: The Duolingo English Test as a Case Study

Abstract (Summary):

This paper, forthcoming in the Handbook for Assessment in the Service of Learning, uses the Duolingo English Test (DET) as a case study to explore the critical role of Responsible AI (RAI) in high-stakes assessment. While AI offers efficiencies in item generation and scoring, it also presents risks, such as bias. The DET addresses these risks through four specific RAI standards—Validity and Reliability, Fairness, Privacy and Security, and Accountability and Transparency—demonstrating how these translate into practical implementation.

The paper details the development of these standards and their connection to broader RAI principles, providing concrete examples of how these practices ensure both test quality (valid score inferences) and equity (fairness for all test takers). The DET serves as a model for how companies can transparently and accountably integrate AI while mitigating ethical risks.

Theme of the Case Study:

- English proficiency test
- AI fairness and unbiased in tests
- (Human in the loop) HiTL and AI
- Responsible AI

Relevance to the Reader:

This paper's relevance to the reader centres on the crucial role of Responsible AI (RAI) in ensuring both the quality and equity of AI-powered assessments. It uses the Duolingo English Test (DET) as a compelling case study, detailing how RAI principles are implemented in practice to mitigate the risks of AI while maximizing its benefits. This is valuable for anyone involved in assessment development, AI ethics, or the use of AI in high-stakes testing. The paper provides a detailed, practical example of how RAI principles can be implemented in a real-world, high-impact setting. This is crucial given the increasing use of AI in education and assessment. It also offers insights into the process of creating an AI-powered assessment, including item generation, scoring, and security measures. It highlights the importance of human-in-the-loop approaches for maintaining quality and equity. AI ethics and governance: The four RAI standards (Validity and Reliability, Fairness, Privacy and Security, Accountability and Transparency) presented serve as a model for organizations developing and deploying AI systems, particularly in high-stakes applications. The alignment with the NIST AI RMF further strengthens the framework's validity. The paper demonstrates how a systematic application of RAI practices can contribute to achieving both test equity (fairness for all test takers) and quality (valid score inferences). This is crucial for ensuring that AI-powered assessments are not biased or discriminatory.



Case Study 7

Gender bias in ads on prominent social media site.

Theme of the Case Study:

- Social Media
- Bias

Abstract (Summary):

Global Witness initiated some paid ads across Facebook in the following countries: UK, the Netherlands, France, India, Ireland, and South Africa. The ads focused on real job vacancies from various fields including pilots, hairdressers and psychologists. The results of who the ads were targeted at through the social media sites algorithm showed a clear gender bias relating to the job vacancy with 91% of individuals being shown the advert for the mechanic vacancy being male and 79% of individuals shown an advert for a pre-school teacher being female.

Upon further research in The Netherlands and France in collaboration with a feminist organisation, it was shown that 97% of individuals shown an advert for a receptionist vacancy were female.

As part of the mandatory selection process when creating an advert, Global Witness selected the objective of targeting those that Facebook thought would be most likely to click on the ad and it was only specified for each ad that it must be shown to adults who live in the country or had recently been in either country.

Relevance to the Reader:

This case study highlights the ingrained bias present in the algorithms that influence the content people see daily, particularly in job advertisements on social media. AI-driven systems, which determine ad placement, often reflect and reinforce existing societal biases, leading to disparities in how job opportunities are presented to different genders.

Social media has evolved beyond a platform for casual interaction; it now serves as a major source of information, support, and career guidance. However, algorithmic bias means that certain job roles may be disproportionately shown to men or women based on historical data rather than individual merit, limiting access to opportunities and reinforcing stereotypes.

Given the reliance of young people on social media for career-related information, it is crucial to raise awareness of this bias and equip them with critical thinking skills to recognise and counteract it. Encouraging transparency in AI systems and advocating for fairer algorithmic practices can help create a more equitable digital job market.

Sources

Case study 1:

- Axios:

<https://www.axios.com/2020/08/19/england-exams-algorithm-grading>

- Wired: <https://www.wired.com/story/skewed-grading-algorithms-fuel-backlash-beyond-classroom>

Case study 2:

- Reuters – "Amazon scraps secret AI recruiting tool that showed bias against women" (2018):

<https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/>

- LinkedIn – "Amazon's Biased Hiring Algorithm - A Wake-up Call for Tech" (2024)

<https://www.linkedin.com/pulse/amazons-biased-hiring-algorithm-2018-wake-up-call-tech-akshay-aryan-dfouc>

- The Independent – "Amazon scraps 'sexist AI' recruitment tool" (2018)

<https://www.independent.co.uk/tech/amazon-ai-sexist-recruitment-tool-algorithm-a8579161.html>

- Leoforce – "Lessons from Amazon's Sexist AI Recruiting Tool" (2024):

<https://leoforce.com/blog/amazon-ai-hiring-bias-case-study/>

- Carnegie Mellon University – "Amazon Scraps Secret AI Recruiting Engine that Showed Biases Against Women" (2018)

<https://www.ml.cmu.edu/news/news-archive/2016-2020/2018/october/amazon-scraps-secret-artificial-intelligence-recruiting-engine-that-showed-biases-against-women.html>

- ACLU – "Why Amazon's Automated Hiring Tool Discriminated Against Women" (2018):

<https://www.aclu.org/news/womens-rights/why-amazons-automated-hiring-tool-discriminated-against>

- IMD – "Amazon's sexist hiring algorithm could still be better than a human" (2018):

<https://www.imd.org/research-knowledge/digital/articles/amazons-sexist-hiring-algorithm-could-still-be-better-than-a-human/>

- Redress Compliance – "Amazon AI Hiring Tool: A Case Study in Algorithmic Bias" (2025):

<https://redresscompliance.com/amazon-ai-hiring-tool-a-case-study-in-algorithmic-bias/>

Case study 3:

https://ceur-ws.org/Vol-3938/Paper_14.pdf

<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1221177/full>

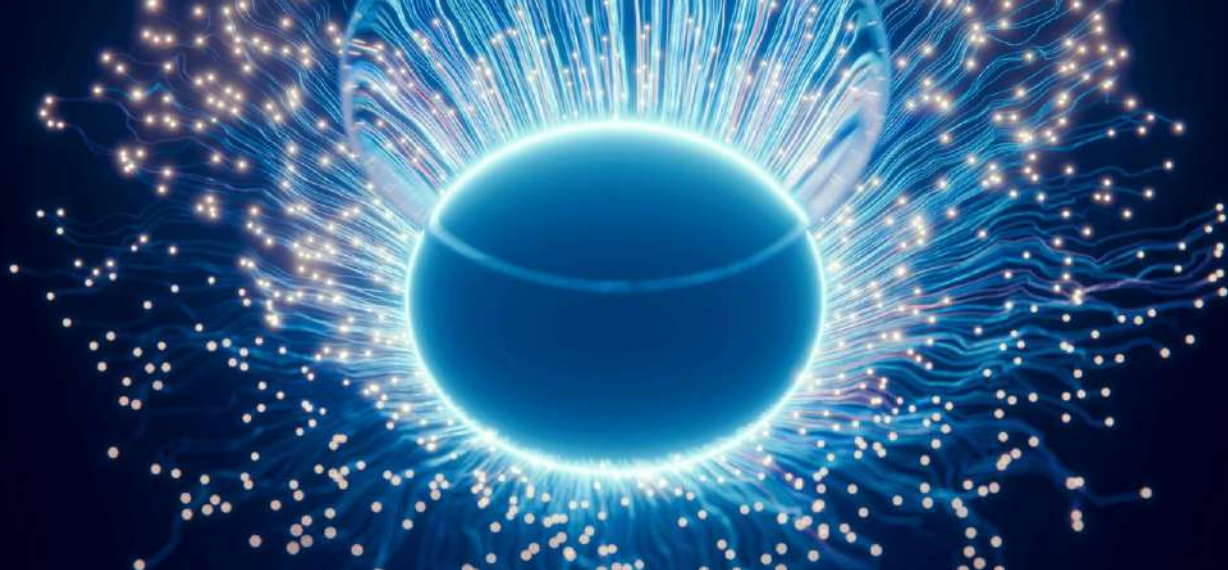
Punoševac, Mina & Nikolić, Anđela. (2024). Transforming Assessment Practices – A Case Study of YouTestMe. CEUR Workshop Proceedings

Case Study 4:

National Youth Assembly on AI Report (2023)

Government of Ireland, Department of Enterprise, Trade and Employment

<https://enterprise.gov.ie/en/publications/publication-files/national-youth-assembly-on-ai-report.pdf>



Sources

Case study 7:

<https://globalwitness.org/en/campaigns/digital-threats/new-evidence-of-facebooks-sexist-algorithm/>

- Published on the 12 June 2023
- Updated on the 13 February 2025

Case study 5:

- Milosevic et al. (2023). Effectiveness of Artificial Intelligence-Based Cyberbullying Interventions from a Youth Perspective
- Dublin City University (DCU) Research, Ireland
<https://doras.dcu.ie/29411/1/milosevic-et-al-2023-effectiveness-of-artificial-intelligence-based-cyberbullying-interventions-from-youth-perspective.pdf>

Case study 6:

- <https://arxiv.org/pdf/2409.07476>
https://www.researchgate.net/publication/383985608_Responsible_AI_for_Test_Equity_and_Quality_The_Duolingo_English_Test_as_a_Case_Study
- Burstein, Jill & LaFlair, Geoffrey & Yancey, Kevin & von Davier, Alina & Dotan, Ravit. (2024). Responsible AI for Test Equity and Quality: The Duolingo English Test as a Case Study.



Conclusion

As we have explored throughout this guidebook, the integration of artificial intelligence throughout the different areas of society; education, employment, media and more, offers both transformative potential and meaningful challenges. The case studies presented demonstrate the diverse and creative ways AI is being harnessed.

What unites these examples is a common thread of innovation driven by the needs and voices of young people. They show that when AI is used thoughtfully and ethically, it can empower young people to take more ownership of their education, employment and personal lives, whilst fostering critical thinking about the technology itself.

However, these advances also underscore the importance of digital literacy, responsible use, and equity of access. Educators, youth workers, and policy-makers all have a role to play in guiding the use of AI in ways that prioritise human connection, ethical practice, and lifelong learning.

This guidebook is not an endpoint but an invitation—to continue exploring, questioning, and co-creating a future where young people are not just users of AI, but active shapers of its role in education and society.

[Valuesai.eu](https://valuesai.eu)

www.valuesai.eu

case study guide 2025 by VALUES is licensed under CC BY 4.0



follow our journey



Co-funded by
the European Union