

## Leon Furze

Reading, Writing, Digital



# Teaching AI Ethics



As we head into the start of Term 1 it's already looking like Artificial Intelligence is going to be one of the most talked about issues in the classroom. Much of the narrative around models like OpenAI's ChatGPT has centred on students using it to cheat on assignments. But I've already been working with schools this year who are much more interested in the potential of these technologies to help, rather than hinder in education.

Earlier this week I wrote a post with some [Practical Strategies for ChatGPT in Education](#). It's proven to be one of the most popular posts I've written on this blog, showing that there's an appetite from teachers wanting to learn how to work with the AI. I also wrote a "[Back to Basics](#)" post for those who had

never heard of ChatGPT, or wanted some grounding in AI and Large Language Models before diving in.

And yet, as much as I enjoy working with the technology, it has many flaws. I think it's our responsibility to discuss the ethical considerations of AI with our students. AI ethics goes beyond the well-documented "algorithmic bias" that results in language models like ChatGPT producing racist and sexist output. In this article, I will explore nine ethical considerations ranging from "beginner" to "intermediate" and "advanced" levels.

## Practical Strategies for ChatGPT in Education

February 16th  
4pm-5:30pm

Eventbrite

Leonfurze.com

[I've turned my previous post – Practical Strategies for ChatGPT in Education – into a live webinar for February the 16th. Click here for details](#)

I've levelled the concerns for two reasons. Firstly, the levels reflect how easy it is to access information and resources on the particular ethical concern, and how likely the concepts are to already fit within your curriculum. For example, "environmental impact" is a beginner-level concern as it has been explored thoroughly in the media, and the climate crisis is already part of many curricula. Secondly, as you move through the levels you and your students will be required to understand and apply increasingly complex concepts and terminology. "Affect recognition" in the advanced level, for

instance, requires some knowledge of facial recognition and the psychology of human emotions.

For each level I've also included a few examples of how and where you could teach students about these issues. Right now, there is no "AI curriculum" for schools. There is no single subject area devoted to Artificial Intelligence. In fact, AI already influences most aspects of our lives, so it is fitting that we should teach AI ethics across *all* of our subject areas. The links embedded throughout this post are there to use as resources, and for going further down the rabbit hole.

**Grab a free PDF infographic of this post:**



[leonfurze.com/aithics](https://leonfurze.com/aithics)

## Table of Contents:

1. [Beginner](#)
  1. [Bias](#)
  2. [Environmental impact](#)
  3. [Academic integrity and "truth"](#)
2. [Intermediate](#)
  1. [Copyright and intellectual property](#)
  2. [Privacy and security](#)
  3. [Data collection and "datafication"](#)

3. [Advanced](#)
  1. [Affect recognition](#)
  2. [Human labour](#)
  3. [Power and hegemony](#)
4. [Ethical Perspectives, Approaches, and Frameworks](#)
  1. [Ethical perspectives](#)
  2. [Ethical guidelines](#)
  3. [“Ethics washing”](#)
5. [Conclusion](#)
  1. [Teaching AI Ethics: Suggested Activity](#)
6. [Resources and Further Reading](#)
  1. [Books](#)
  2. [Resources](#)
  3. [AI ethics policies and guidelines](#)

# Beginner

## Bias

Artificial Intelligence comes in many forms, but all require data. ChatGPT, for example, is a large language model that is trained on a huge dataset which includes the [“common crawl”](#); a text-based database of over 12 years’ worth of web pages. These datasets give the models tremendous capabilities, but they are also inherently [biased](#). Indiscriminately [“scraping” the internet lets in the bad along with the good](#), meaning that the dataset can contain racist, sexist, ableist, and otherwise discriminatory language.

Unfortunately, AI large language models hold up a mirror to internet society, and the reflection isn’t pretty. Like other societies, the online community underrepresents marginalised groups, and overrepresents others. [The prevalence of racism and bigotry on sources like Reddit and Twitter](#) can bleed through the datasets and be reproduced in the output.

Bias can also come from the methods of training and reinforcement used when developing the AI systems. For example, [police in the US have used systems for “predictive policing”](#) which use algorithms to predict people likely to commit crimes. These algorithms disproportionately target poor, Black, and Latinx communities, reinforcing existing systemic biases.

## Teaching points

- “Garbage in, garbage out” is a common phrase in computing. Unfortunately, much of the data that goes into AI models is biased, resulting in discriminatory output.
- Datasets can exclude minority groups or marginalised people.
- AI algorithms can perpetuate systemic biases, for example by targeting communities who already suffer under biased judicial systems.

## Subject examples

- Humanities (Legal studies): What legal precedents exist for protecting marginalised groups from discrimination? Is a biased algorithm legal?
- English and Literature: How have certain groups been silenced or oppressed throughout history? What is the implication of this “gap” in the written record of the internet when it is used as data to train an AI?
- Mathematics: What is an algorithm? How do algorithms and probability link to policing and other societal functions?

# Environmental impact

The technology industry as a whole has an enormous impact on the environment. Most devices from smartphones to laptops incorporate metals like lithium and rare earth minerals which are in short supply and costly to

extract. The mining and refining of these products adds to the environmental impact of developing the technologies that AI is built from. These costs [include soil erosion, water pollution, and greenhouse gas emissions.](#)

AI computing is increasingly carried out in “the cloud”. Cloud services sound like an ethereal and temporary arrangement, but the name actually hides the physical reality of the technology. [Cloud computing relies on huge data centres and infrastructure, all of which consumes energy and produces waste.](#)

[Although many of the major companies such as Google, Microsoft, and Meta have pledged to make their data centres carbon neutral,](#) in reality this often means engaging in carbon-trading or offsetting schemes rather than actually reducing the amount of waste or environmental damage. This “greenwashing” is heavily criticised by those who would rather see an actual reduction in emissions.

## Teaching points

- The infrastructure behind AI – including the devices it runs on – has a huge impact on the environment through processes like mining and industry.
- Cloud computing still requires enormous amounts of energy.
- AI companies have been accused of “greenwashing” when they commit to reducing their carbon footprint, but do not actively reduce their impact on the environment.

## Subject examples

- Humanities (Geography): What is the impact of the climate crisis on different global populations? Which parts of society are responsible for the most emissions due to AI technologies?

- Science: Why does computing use so much energy and produce so much waste heat? What is the impact on the environment?
- Design and Technology: How might we design more sustainable systems for the manufacturing of AI technologies?



*The cloud isn't as fluffy as it sounds. Via Midjourney. Prompt in alt text.*

## Academic integrity and “truth”

Academic integrity – or using AI language models to cheat – has been by far the biggest potential issue of AI covered recently in the media. [There have been widespread fears that students will use language models like ChatGPT to write essays, answer questions, and cheat on assignments.](#) These fears seem particularly strong in secondary and tertiary education, where many assignments are provided in written form.

It is also still unclear to what extent using an AI constitutes “cheating”. It is not, strictly speaking, plagiarism as the output of the model is not copied from another source. Rather, the output is an original creation which is generated “[probabilistically](#)”. Knowing where to draw the line raises ethical questions about academic integrity and honesty. This has already led some universities to permit the technologies as long as they are credited. In fact, this article was written with the assistance of ChatGPT. All of the words you’re reading are mine (I happen to enjoy writing), but I used the AI to help organise the structure and to fill some of my knowledge gaps in different subject areas. I’ll explain the process in full in my next post.

As well as cheating, there are concerns that AI will be used to [produce massive amounts of “fake news” or deliberately harmful media](#). This may be unintentional – one of the biggest current flaws of most language models is that they can generate very convincing lies. Or, people may use these technologies maliciously to spread political misinformation or otherwise cause harm.

## Teaching points

- Academic integrity is an important part of any education system. Think about the rules and regulations that exist, and why they are in place.
- AI models produce original output that cannot be detected by plagiarism detectors (and “AI detectors” are currently hit and miss).
- AI can be used to produce “fake news” or deliberately misleading information.

## Subject examples

- English: If an AI can write an essay, what is the point of writing essays? how do essays help us to build knowledge and not just demonstrate it?

- Religion: What are the ethical and moral implications of academic integrity?
- Humanities (Legal studies): Is using an AI to write an essay cheating? Australia has laws against contract cheating (getting another person to produce an essay on your behalf). Is using an AI a form of contract cheating and therefore illegal?

# Intermediate

## Copyright and intellectual property

Closely related to truth and integrity are ideas of intellectual property and the legal concerns of copyright. Copyright issues have been particularly prevalent in AI image generation. AI image generators like Stable Diffusion, DALL-E 2, and Midjourney are trained on images “scraped” from the internet. [These images are then broken down and run through the AI algorithms, so that they can later be reconstructed.](#)

This has resulted in artists’ “styles” being used in AI image generation without their permission. Many artists believe that this infringes their [intellectual property rights](#), and is an ethical issue. An often-used counterargument is that all art is based on other artists’ work, and therefore the machine is simply replicating those processes. Class action lawsuits have already been filed against some AI image generators on behalf of artists.

Large language models like ChatGPT also incorporate huge amounts of other writers’ work. Where writing is publicly available – such as out of copyright books or CC journalism and articles – it can be incorporated into the dataset. Even when writing is protected by copyright it can become part of the datasets. Prompting a language model to write something *in the style of* another author could be viewed in the same way as an image generator adopting another artist’s style.

[There are also question marks over who owns the copyright to materials produced by AIs such as image generators and language models.](#)

## Teaching points

- AI language models and image generators may breach copyright by appropriating others' work.
- These models give users the ability to write and produce art in the style of other authors and artists without the original creator's permission.
- Laws around these technologies are still murky, but there are developments happening all the time.

## Subject examples

- Visual Arts: Is it art? If a user can generate an image in the style of another artist with just a few prompt words, does the digital output count as "real art"?
- Performing Arts: Some AI models can produce music and lyrics as well as visual art. Is it possible to create a complete AI performing artist? If we can, does it mean we should?
- English: Does producing a piece of writing in the style of another author infringe their intellectual property rights?



*I call this one "Melbourne skyline in the style of picasso and van gogh –ar 3:2 –q 2 –v 4" but is it art? And is it legal? Via Midjourney.*

## Privacy and security

Privacy is a major concern in the development and use of AI systems. As these technologies become more sophisticated and integrated into our lives, there are increasing concerns about the [collection and use of personal data](#), [data breaches](#), and the [lack of transparency](#) in AI decision-making.

One of the most prominent examples of these issues can be found in the use of facial recognition technology. This technology, which is used in a variety of applications such as security, surveillance, and marketing, has been criticised for its potential to violate individuals' privacy and civil rights. For example, [facial recognition systems have been known to have higher error](#)

[rates for people with darker skin tones](#), and have been used to target and monitor marginalised communities as discussed earlier in “bias”.

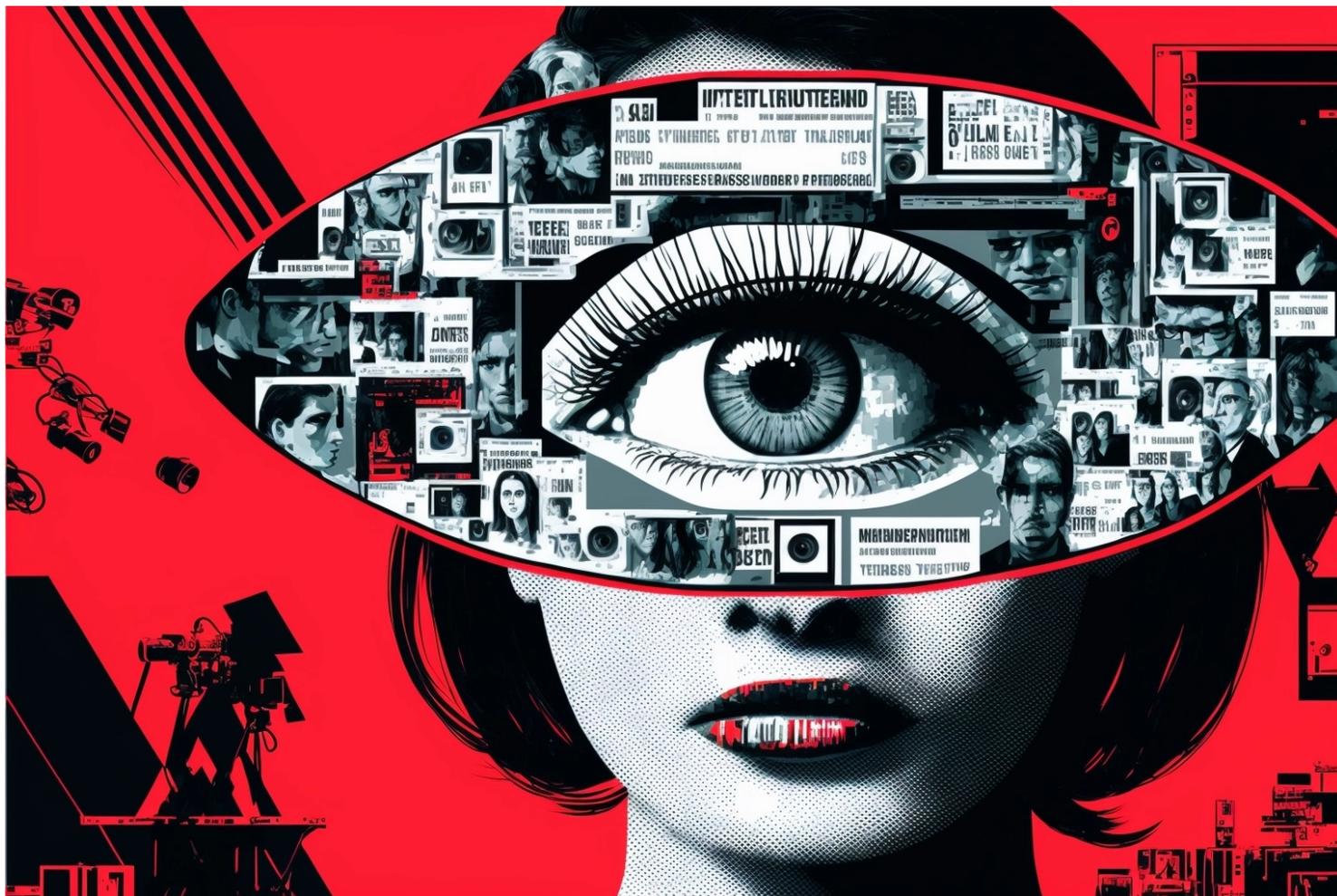
Another example of privacy concerns with AI systems is targeted advertising. AI-powered algorithms are used to analyse data on individuals' online activities in order to deliver targeted ads. Whilst this may seem harmless, it raises [concerns about data privacy](#), data breaches, and the use of personal data for commercial gain.

### Teaching points:

- AI systems raise concerns about the collection and use of personal data, data breaches, and the lack of transparency in decision-making.
- Facial recognition technology has been criticized for its potential to violate individuals' privacy and civil rights, and for its higher error rates for people with darker skin tones.
- Targeted advertising raises concerns about data privacy and the use of personal data for commercial gain.

### Subject examples:

- Humanities (Legal Studies): What is the impact of AI on data protection laws such as GDPR and the protection of personal data?
- Mathematics: How can we analyse the data sets used by AI systems for potential biases and privacy issues?
- Health and Physical Education: What are the privacy concerns surrounding personal health data and its use in AI-powered healthcare technologies?



*Every click and like goes towards powering AI surveillance. Image via Midjourney. Prompt in alt text.*

## Data collection and “datafication”

The phrase “[data is the new oil](#)” crops up everywhere when you start researching AI. As I wrote about earlier in “bias”, Artificial Intelligence is powered by huge amounts of data. The oil analogy suggests both data as fuel, but also the costly, dangerous, and extractive process of data-mining. In the constant quest for more and more data, the companies that develop AI systems sometimes revert to unethical practices.

“[Datafication](#)” is a term used for turning all parts of our life into a data point to be fed into an AI algorithm. As per the privacy discussion above, this should raise some serious concerns. From location data to health data,

shopping habits, likes, clicks, and views, almost every interaction we have with technology is fed into an algorithm somewhere.

As we become commodities, we open ourselves up to exploitation. One major ethical concern with “datafication” is that fact that the users become the products, and that the free-labour of the users is used to generate capital for the platform owners.

“Big Data” also contributes to many of the issues we have described so far, including bias and [discrimination](#). Any data collected by the devices we wear and use or the platforms we subscribe to ultimately becomes part of the algorithm’s “world view”. Unfortunately, because not everyone in the world has access to these technologies, that worldview is by definition missing some very important data.

### Teaching points

- Data is the new oil, both in the sense of being fuel, and in that it is costly and damaging to extract.
- “Datafication” is the process of turning every aspect of our lives into data.
- Users become products, and user data becomes capital.
- Big Data doesn’t include marginalised groups, and therefore doesn’t truly represent society.

### Subject examples

- Humanities (History): What is the historical and societal context of data collection and the impact of datafication on different communities?
- Design and Technology (Digital Technology): How can we design and develop ethical data collection practices and data privacy measures for AI systems?

- English: Based on recent media, how might we critically analyse datafication and its implications on privacy and data protection?

Enter your email



Join the mailing list

# Advanced

## Affect recognition

Affect recognition means interpreting a person's emotions through their facial expressions, body language, speech patterns and actions. It's a controversial practice that has been widely [criticised for poor research methodologies and inconsistent results](#).

Despite these controversies, affect recognition is an industry [worth billions of dollars](#). It is also an industry that has already made its way into education. A system named 4 Little Trees, developed in Hong Kong, [claimed to be able to monitor children's facial expressions](#) and to assign labels for emotions such as 'happy', 'sad', and 'angry'. The system also claimed to be able to identify motivation and to predict grades.

Affect recognition is problematic on a number of levels. As well as the aforementioned question mark over its accuracy, many people question whether emotions should be "datafied" at all. There are privacy concerns with affect recognition being built into surveillance technology, including in schools. And, similar to the issues with bias discussed earlier, affect recognition technology can perpetuate discrimination. In one example, an algorithm trained to identify possible "terrorist behaviour" [resulted in racial profiling](#).

## Teaching points

- Affect recognition is the detection of emotions through facial expression, actions, tone of voice, and other data.
- The science is unreliable, but the industry is worth billions.
- There are many ethical concerns including surveillance, privacy, and discrimination.

## Human labour

The ethical concern of AI and human labour is a two sided coin. On the one hand there are always fears that [machine automation will replace jobs](#), even in white collar industries like law and finance. On the other is the fact that current AI systems actually rely on a tremendous amount of dangerous, low-paid human labour.

The “robots taking our jobs” argument goes back a long way. [In the 16th Century, Queen Elizabeth I rejected an application for a patent on a stocking making machine](#), for fears it would put too many stocking-makers out of work. In more recent years, old fears of AI replacing human “knowledge work” have been reignited by ever-more powerful models like GPT-3. And though most commentators are quick to claim that AI will never replace teachers, [some have made predictions](#) that some or all parts of the job could be automated by as early as 2027.

Hidden beneath the rhetoric of the jobs AI will destroy, however, is an unseen narrative of the jobs it currently requires to function. It is useful for the companies behind AI technology that the public views it as something mysterious and almost magical. Current advances like ChatGPT and Midjourney seem to be able to produce countless outputs in text and image with little input. But there is human labour powering the magic.

[A recent article by Time magazine](#) explored the harsh conditions of the Kenyan workers employed by OpenAI to label inappropriate data for its language model. Working for less than \$2 an hour, these labourers were partly responsible for training an AI algorithm to identify harsh language, graphic, sexual, and violence phrases, and other “toxic” text. Workers were required to read and label huge amounts of this data, with some reporting the experience as deeply traumatic.

### Teaching points

- AI could replace jobs, even in traditionally “white collar” industries.
- There are low-paid, risky jobs currently used to train AI models.
- AI might replace some jobs, but we also need to be mindful of the human labour cost that goes into its production along the way.

### Subject examples

- Humanities (History): What is the history of the “robots taking our jobs” argument and how has it evolved over time?
- Humanities (Economics): What is the potential impact of AI on employment and the labour market?
- Science (Psychology): What is the psychological of low-paid labor on workers and the potential for trauma?



*The human costs of AI labour are more than just job cuts.  
Image via Midjourney. Prompt in alt text.*

## Power and hegemony

This final ethical concern brings us full circle back to “bias”, but with a more nuanced perspective. Because the data AI models are built on is “frozen in time”, it represents a static world view which encodes existing power and hierarchies in society. The reinforcement of the [hegemony](#) can further oppress and marginalise already disadvantaged people.

Think of AI as a self-perpetuating cycle. The datasets encode a certain power structure into the model – often the dominance of a heterosexual, white, Western, male perspective due to the volume of content on the internet from that lens. This is then reflected in the output, which may be used to train future models by generating “synthetic data”. Although [efforts](#)

[are underway to make “fair” synthetic data](#), it has still be found to reproduce biases.

AI also reinforces global hegemonies both in political and corporate terms. Countries and organisations need access to wealth, energy, and resources to successfully train and scale up AI models. This means that powerful AI is increasingly concentrated in the hands of those who already have the most. Actions like those outlined above in “human labour” further entrench the divide between the wealthy countries who produce AI and the poorer countries who bear the brunt of the human and environmental costs.

### Teaching points

- AI has a worldview which includes encoded biases and perspectives.
- The replication of these biases reinforces existing hegemonies and power structures.
- AI concentrates wealth in the hands of the already wealthy.

### Subject examples

- Humanities (Geography): Geography: Explore the global distribution of wealth, energy, and resources in relation to AI development and how it impacts different countries and regions.
- Mathematics: Explore the statistical analysis of AI-generated data, including the detection and measurement of biases and the impact on decision making.
- English and Literature: How is language examined representation in AI-generated text and speech. How can a Marxist critical perspective illuminate some of the problems of AI and power?

# Ethical Perspectives, Approaches, and

# Frameworks

In discussing these ethical issues with students, it can be helpful to bear in mind some existing ethical perspectives, frameworks, and approaches. In this section I'll explore some general ethics as well as some frameworks for ethical and "Responsible" AI produced by companies like Google and Microsoft.

Apply these frameworks and perspectives to your discussions with students, for example by providing them with the examples and asking them to produce their own ethical frameworks.

## Ethical perspectives

There are many ethical perspectives which could be applied to AI. [In this paper](#), for example, the author discusses [Terry Bynum's](#) principles of virtue ethics and human flourishing, translated from Aristotle into modern times. I've used ChatGPT to "translate" one step further by asking for a simplification that may be more suitable for using with students. Bynum's original is on the left.

1. Human flourishing is central to ethics.
2. Humans as social animals can only flourish in society.
3. Flourishing requires humans to do what we are especially equipped to do.
4. We need to acquire genuine knowledge via theoretical reasoning and then act autonomously and justly via practical reasoning in order to

1. Being happy and healthy is important for doing the right thing
2. People need other people to be happy and healthy
3. To be happy and healthy, people need to do what they are good at
4. To know what is right, people need to learn and think for themselves
5. Making good choices is key to doing what is right.

flourish.

5. The key to excellent practical reasoning and hence to being ethical is the ability to deliberate about one's goals and choose a wise course of action.

Discussions of “human flourishing” and basic virtues then feed into more specific ethical guidelines developed by organisations and countries to help govern the creation of AI.

## Ethical guidelines

Here are several examples of guiding principles used in AI ethics or “responsible AI”. There are many more available, some of which I will link in the resources section of this post.

### **EU High-Level Expert Group on Artificial Intelligence**

1. Human agency and oversight
2. Technical robustness and safety
3. Privacy and data governance
4. Transparency
5. Diversity, non-discrimination and fairness
6. Social and environmental wellbeing
7. Accountability

### **Google**

1. Be socially beneficial
2. Avoid creating or reinforcing unfair bias
3. Be built and tested for safety

4. Be accountable to people
5. Incorporate privacy design principles
6. Uphold high standards of scientific excellence
7. Be made available for uses that accord with these principles.

## **Industry.gov.au Australian Ethics Principles**

1. Human-centred values
2. Fairness
3. Privacy protection and security
4. Reliability and safety
5. Transparency and explainability
6. Contestability
7. Accountability

There are some key overlapping areas between these guidelines, and some which are unique. In teaching AI ethics, explore many different examples and discuss with students which are more appropriate, and why.

## **“Ethics washing”**

One final point to note is that just because organisations develop AI Ethics or Responsible AI principles, doesn't mean they follow them. Just like the “greenwashing” of environmental concerns discussed earlier, many AI companies have been accused of paying lip-service to ethical concerns.

AI Ethics principles are often non-binding and cannot be enforced by law. To do that, we need laws and regulations imposed by states and countries.



*"this my hand will rather the multitudinous seas incarnadine".*

*Can Responsible AI wash away the sins of big tech? Image via Midjourney. Prompt in alt text.*

## Conclusion

We are at a watershed moment in terms of Artificial Intelligence like ChatGPT. Although the technology has existed for years, the release of OpenAI's chatbot has thrust it into the limelight, raising awareness for both teachers and students.

Beyond the narrative of cheating and dramatic claims of the end of industrial-era education, AI offers real opportunities to personalise learning and help with differentiation, diversity, and inclusion. But there is a shadow side to AI which cannot be ignored.

The ethical concerns of AI are complex – I haven't even scratched the surface in this article, but hopefully I have provided enough for teachers to start exploring down their own rabbit holes. I'll end with a few final suggestions for teaching AI ethics, and resources and further reading. If you've enjoyed this article, please share it.

## Teaching AI Ethics: Suggested Activity

This activity comes via ChatGPT with some back-and-forth to refine the idea. I'll explore the method behind creating resources like this in more detail in a future post. For now, here are the prompts and the activity:

**Prompt: Suggest a teaching idea for these ethical issues. Do not align to a particular subject area. Build out robust and interesting lesson activity ideas which can be applied to any of the 9 areas. Do not align activities with the 9 areas. Use contemporary teaching practices, and a mix of discussion, research, student centred, and explicit instruction. Produce a title for the activity (use markdown to format headings), one or two learning intentions, and the description of the activity. Limit activity to a maximum of 50 minutes.**

### Activity Title: "Ethical Decision Making in AI"

- **Learning Intention 1:** Students will understand the importance of ethical decision making in the development and implementation of AI technology.
- **Learning Intention 2:** Students will be able to identify and analyse potential ethical issues in a given AI scenario.

### Activity Description:

1. Begin the lesson by introducing the concept of ethical decision making in AI and its importance.
2. Divide the class into small groups and provide each group with a different AI scenario (e.g. facial recognition technology in law enforcement, self-driving cars, AI-powered virtual assistants).
3. Have each group brainstorm potential ethical issues that may arise from the implementation of the AI technology in their scenario.
4. Each group will then present their findings to the class and engage in a class discussion on the potential ethical issues identified by each group.
5. As a class, generate a list of best practices for ethical decision making in AI development and implementation.
6. As an extension, students can research real-life cases where ethical issues have arisen in AI and present it to the class.

**Prompt: Focus on the activity “Ethical Decision Making in AI”. Create a handout resource for the different scenarios for step 2. Produce 6 scenarios each with a subheading (e.g., facial recognition in law enforcement, self-driving cars, AI-powered assistants, AI powered tutors). The scenario must include a brief description of the scenario and the pros and cons and ethical concerns.**

**Handout:**



Teaching AI Ethics – Resource via leonfurze.com / ChatGPT

**1. Facial Recognition in Law Enforcement**

- Description: The use of facial recognition technology in law enforcement to identify and track suspects and potential criminals.
- Pros: The technology can help to quickly identify and track criminals, potentially reducing crime rates. It can also help to improve the efficiency of police investigations.
- Cons: There are concerns that the technology may be inaccurate, leading to false identifications and wrongful arrests. There are also concerns about the technology's impact on civil liberties and privacy. Ethical concerns include issues of bias and discrimination, particularly against marginalised communities.

**2. Self-Driving Cars**

- Description: The use of AI technology to power autonomous vehicles.
- Pros: Self-driving cars have the potential to reduce accidents caused by human error, as well as traffic congestion. They could also improve mobility for those who are unable to drive due to disability.
- Cons: There are concerns about the reliability of the technology, particularly in the event of failures or unexpected situations. There are also concerns about the impact of self-driving cars on employment, particularly in the transport sector. Ethical concerns include issues of liability for accidents and issues of access to the technology for marginalised communities.

**3. AI-Powered Tutors**

- Description: The use of AI technology to power educational programs that can provide personalised instruction and feedback to students.
- Pros: AI-powered tutors can provide personalised instruction and feedback to students, potentially improving educational outcomes.
- Cons: There are concerns about the reliability and accuracy of the AI technology, particularly in understanding and responding to students' individual needs. There are also concerns about the impact of the technology on employment in the education sector. Ethical concerns include issues of access to the technology for marginalised communities and issues of bias and discrimination in the AI's assessment of student performance.

**4. AI-Powered Assistants**

- Description: The use of AI technology to power virtual assistants, such as those found on smartphones and smart home devices.
- Pros: AI-powered assistants can improve productivity and efficiency in daily tasks, such as scheduling and information retrieval.
- Cons: There are concerns about the privacy of having a device constantly listening in on conversations and collecting data. There are also concerns about the reliability of the technology, particularly in understanding and responding to complex requests. Ethical concerns include issues of discrimination, particularly in the areas of voice recognition.

<https://leonfurze.com/2023/01/26/teaching-ai-ethics/>

Teaching AI Ethics – Resource via leonfurze.com / ChatGPT

**5. AI-Powered Healthcare**

- Description: The use of AI technology to assist in the diagnosis and treatment of patients in the healthcare sector.

**6. AI-Powered Finance**

- Description: The use of AI technology to assist in financial decision making, such as investment management.

[handout](#)

# Practical Strategies for ChatGPT in Education

February 16th

4pm-5:30pm

**Eventbrite**

Leonfurze.com

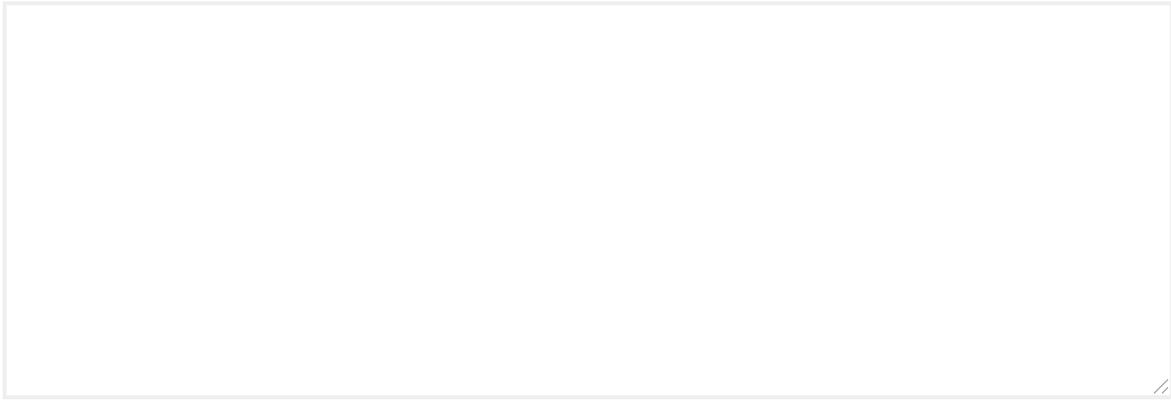
[I've turned my previous post – Practical Strategies for ChatGPT in Education – into a live webinar for February the 16th. Click here for details](#)

If you have questions, comments, or would like to chat about teaching with AI, use the form below to get in touch:

**Name** (required)

**Email** (required)

**Message**



Contact Leon



*Towards an ethical future for AI in education. Image via Midjourney. Prompt in alt text.*

# Resources and Further Reading

## Books

- Kate Crawford – Atlas of AI
- Virginia Eubanks – Automating Inequality
- Shoshana Zuboff – The Age of Surveillance Capitalism
- James Barrat – Our Final Invention
- Nick Srnicek – Platform Capitalism
- Mark Andrejevic – Automated Media

## Resources

[Ethical guidelines on the use of AI in education – European Council](#)

[Teaching kids the ethics of AI – MIT media lab](#)

[AI in education – ISTE](#)

## AI ethics policies and guidelines

- [Microsoft Responsible AI Principles](#)
- [Google AI Principles](#)
- [Industry.gov.au Australian AI Ethics Principles](#)
- [DeepMind Safety and Ethics](#)
- [OpenAI Charter](#)
- [Stability.ai FAQ](#)
- [Midjourney](#)

Enter your email



Join the mailing list

Leon Furze

[January 26, 2023](#)

[AI](#)

[AGI](#), [AI](#), [arts](#), [autocomplete](#), [bias](#), [chatbot](#), [ChatGPT](#), [classroom](#), [Common Crawl](#), [computer science](#), [dataset](#), [economics](#), [education](#), [ethical considerations](#), [ethics](#), [expert systems](#), [fine-tuning](#), [GPT-3](#), [history](#), [Language model](#), [large language model](#), [literature](#), [machine learning](#), [mathematics](#), [maths](#), [narrow AI](#), [Natural language processing](#), [NLP](#), [OpenAI](#), [output](#), [PE](#), [philosophy](#), [power](#), [prediction](#), [predictive text](#), [religion](#), [science](#), [strong AI](#), [teachers](#), [technology](#), [training](#), [user interface](#), [weak AI](#)

## 2 responses to “Teaching AI Ethics”

[Utilising AI and ChatGPT as Educators \(Lesson Plans, Feedback, Activities, Communication, etc.\) – AlandAutomation](#)

[February 5, 2023 at 2:09 am](#)

[...] For more information, check out the following articles by Leon Furze:

Article 1 – <https://leonfurze.com/2023/02/01/advanced-techniques-for-chatgpt-in-education/> Article 2 –

<https://leonfurze.com/2023/01/23/practical-strategies-for-chatgpt-in-education/> Article 3 –

<https://leonfurze.com/2023/01/26/teaching-ai-ethics/>

[...]

Loading...

[Reply](#)

[A Brief History of Artificial Intelligence – Leon Furze](#)

[February 11, 2023 at 9:00 am](#)

[...] sure that we don't get swept up in another cycle of AI hype and lose sight of the very real ethical and social concerns of these [...]

Loading...

[Reply](#).

## Leave a Reply

Enter your comment here...

© Leon Furze 2022

[Blog at WordPress.com.](#)