

Received: 17 July 2025 | Revised: 28 November 2025 | Accepted: 28 November 2025

# On ethical AI principles

Stephen Downes<sup>1</sup> 

<sup>1</sup>Casselman, Ontario, Canada

Correspondence:  
Stephen Downes | e-mail: [stephen@downes.ca](mailto:stephen@downes.ca)

---

## Abstract

Artificial Intelligence (AI) is a set of digital tools that can perform functions traditionally limited to human capability, for example, reviewing, summarizing, translating, and composing. In the last few years, there has been widespread discussion and experimentation for its application in education. Consequently, a contentious ethical debate has arisen around its appropriateness for educational functions and how such functionality may be ethically applied. Some argue that the ethical use of artificial intelligence in education can be defined through the use of a set of commonly held principles, such as sustainability, accountability, and fairness. This article challenges that presumption. A list of nine such principles is offered. Each principle is considered in detail and analysed to identify underlying assumptions, points of conflict, and other ethical considerations. It is argued that much of what is offered as a set of ethical considerations reflects, in fact, a political argument and perspective. There is thus no set of ethical principles that can be regarded as a consensus opinion on the ethics of the use of artificial intelligence in education.

---

## Keywords

Ethics; principles; AI; education; fairness



## 1 Introduction

The use of artificial intelligence to support learning has long been an ambition of educational technologists and the sector in general. Readers will be familiar with the trope book machine (Figure 1), Pressey's (1926) teaching machine, Skinner's (1968) teaching machine, and the World's Fair "autotutor" (Novak, 1964).

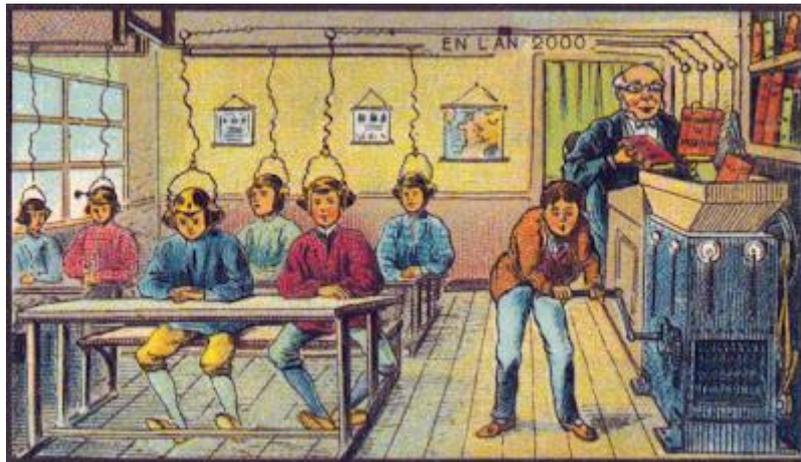


Figure 1: Postcard from the World's Fair in Paris -- Circa 1889 A Futuristic Image of Learning. Image: Wikimedia Commons via McRae and Bower (2013).

In more recent years, the use of artificial intelligence has turned to drawing on statistical inferences to evaluate or predict learning outcomes. The domain of "learning analytics", for example, is "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" (SOLAR, 2025). Learning analytics began as an application of data mining and machine learning (Baker & Inventado, 2016), but by the 2020s was employing neural network technology (Sghir et al., 2023).

Even before the launch of ChatGPT in 2022, concerns were expressed over the use of learning analytics and artificial intelligence to support learning. Hakimi et al. (2021) conducted a systematic qualitative analysis of research in the ethics of digital trace data use in education and as Ye (2022) writes, "found that privacy, informed consent, and data ownership are the most commonly and extensively discussed ethical issues" (p. 613). In 2017, the University of Michigan "Learning Analytics Guiding Principles" cited respect, transparency, accountability, empowerment, and continuous consideration (Regents of the U of M, 2025).

These arguments – and there were many more – are in accord with what appears to be a broad consensus around the idea that the application of AI in education, if it is to be allowed at all, must be subject to ethical consideration, and that as well, these ethical considerations can be addressed as a set of principles coalescing around the ideas that AI should cause no harm, respect individual rights, and accountability (Floridi et al., 2018).

The intent of this article is to challenge that consensus. It is argued that what people have been describing as 'ethical AI principles' actually represents a specific political agenda, and not an ethical agenda at all. While the principles appear to describe an ethical stance, when their clarity and coherence are challenged, it becomes clear that they obscure, rather than enlighten, the actual decisions being made. This argument is illustrated through an outline of some of them.

## 2 The ethical principles, listed

### 2.1 Major ethical AI principles

Here's how Google's search AI<sup>1</sup> describes some major principles of ethical AI; they are drawn from a wide variety of documents from various agencies that say much the same thing. Your Google Search mileage may vary. I quote:

Ethical AI principles focus on ensuring AI is developed and used in ways that are beneficial and respectful of human values and rights. Key areas include fairness, transparency, accountability, privacy, security, and inclusiveness. Here's a more detailed look at some key principles.

*Fairness:* AI systems should not discriminate against individuals or groups. This includes ensuring that training data is representative and that models are not biased.

*Transparency and Explainability:* AI systems should be designed in a way that allows users to understand how they work. This can help build trust and identify potential biases or errors.

*Accountability:* There should be clear lines of responsibility for the development, deployment, and use of AI systems. This helps ensure that AI is used responsibly and that those responsible can be held accountable for any negative consequences.

*Privacy and Data Protection:* AI systems should protect individuals' personal data and respect their privacy. This includes ensuring that data is collected and used ethically and that individuals have control over their data.

*Inclusiveness:* AI systems should be designed to be accessible to all users, regardless of their background or abilities. This can help ensure that AI is used to benefit all members of society.

*Reliability and Safety:* AI systems should be reliable and safe to use. This includes ensuring that they are not vulnerable to errors or malicious attacks.

*Human-centered design:* AI systems should be designed with the needs of humans in mind. This includes ensuring that AI systems are easy to use, understandable, and beneficial.

*Non-maleficence:* AI systems should be designed to avoid harming individuals, society, or the environment. This includes considering the potential risks and consequences of AI systems before they are deployed.

*Sustainability:* AI systems should be designed to be sustainable and environmentally responsible. This includes considering the energy consumption and other environmental impacts of AI systems.

These principles are not exhaustive, but they provide a starting point for developing and deploying AI in an ethical and responsible manner.

### 2.2 The relevance of ethical AI principles to education

The ethical principles as stated here are directly applicable to the employment of AI in education. Numerous works have described this employment, for example, the U.S. Office of Educational Technology (Cardona et al., 2023). AI may support adaptive learning to “meet students where

---

<sup>1</sup> This is the AI-generated results produced when using Google Search. It is not referenced by any particular name. Hence the name, “Google’s search AI”.

they are” (Cardona et al., 2023, p. 18). It can support social and other aspects of learning, neurodiverse learners, creative tasks, and wider educational goals (Cardona et al., 2023, p. 20).

The authors list a set of “open questions” that explicitly reference ethical principles such as the ones just listed, for example, “When AI is used, are students’ privacy and data protected” and “How strong are the processes or systems for monitoring student use of AI for barriers, bias, or other undesirable consequences?” (Cardona et al., 2023, p. 23).

Similarly, the authors caution that the application of AI to support teaching ought to be constrained by these principles. Concerns listed (see Figure 7 in Cardona et al., 2023, p. 27) include algorithmic transparency, user control of data, engaging stakeholders who are diverse, and evaluation for bias. Human-centered design is explicitly endorsed, “keeping a humanistic view of teaching front and center” (Cardona et al., 2023, p. 25). One of the report’s “key recommendations” is “inspectable, explainable, overridable AI” (Cardona et al., 2023, p. 34).

The application of AI to educational assessment is similarly constrained by these principles. AI can enhance feedback loops, offer automated essay scoring, and help with formative assessment (Cardona et al., 2023, pp. 39-41). Such applications must be monitored to ensure they are fair and free from bias, the authors argue (Cardona et al., 2023, p. 43).

Therefore, we can see that AI has a number of potential uses in education, that these are being considered at the highest level, and that the ethics of such uses are being discussed with reference to the list of principles described in section 2.1.

What will be argued in the sections below is that such discussions do not represent an ethical consensus. That does not mean the proponents, such as the U.S. Office of Educational Technology, are wrong to take such factors into consideration. Obviously, they should. Such consideration, however, must go beyond the assumption of unanimity of support for the principle in question. Many such “ethical” principles are also social, cultural, and political in nature, and there is a wide variance of opinion around the world on their expression and their applicability.

### 3 The ethical principles, challenged

#### 3.1 Fairness

Fairness is the idea that AI systems should not discriminate against individuals or groups. This includes ensuring that training data is representative and that models are not biased. Thousands of papers have addressed this issue (Kheya et al., 2024).

Is fairness an ethical principle, though? Ever since John Rawls's landmark *A Theory of Justice* (Rawls, 1971), fairness has been considered one of the bedrock principles of liberal democracy. The spectacle of Donald Trump claiming repeatedly about how “unfairly” he has been treated illustrates the hollowness of the concept, however. We are often told “life isn't fair” (Pemberton, 2022), and that truism seems to be more universal than the principle of “justice as fairness”.

In fact, there are many cases where we do not want fairness to prevail. Some examples are so obvious they are absurd: we don't want wild animals, diseases, and natural disasters to have “a fair go” when it comes to taking human lives. As fun as it would be to see hunters required to engage in “a fair fight” with the bear, we just don't allow this.

Similarly, we do not believe that noxious individuals and ideologies should have a fair chance of success. With the exception of Fox News, nobody believes we should be fair to fascism. A mass killer like Clifford Olson may receive a fair trial, but only after a very unfair campaign to capture him. Cannibalism isn't presented fairly as a lifestyle choice in classrooms.

There are also many cases where we don't care whether fairness applies or not. In sports, for example, we give awards to the highest, fastest, and strongest, which isn't very fair. In a fair competition, arguably, these differences in ability would be handicapped; that's how it's done in golf and in bowling; why not the Olympics? It's not fair that some people are born in Somalia and others are born in Switzerland, yet we have a global system of laws ensuring that system of unfairness is not corrected. Who decides when fairness applies and when it doesn't?

It is moreover clear that there is nothing like a universal commitment to fairness. It is certainly observable that those who are in power and authority frequently manipulate the rules to preserve their position in society. There has been, for example, a recent negative response to the principle of equity (Drenon, 2025), even though it is widely known that various marginal groups are subjected to systemic discrimination. The ethic of "protecting your own" (whatever that amounts to) is widespread and widely practiced and often runs counter to fairness (for example, Hennessey, 2025). We may not agree with it, but its existence in society is a fact.

As the Belmont Report notes "the sense of 'fairness in distribution' or 'what is deserved'" (DHEW, 1978, p. 8) can be viewed from numerous perspectives, each of which needs to be considered, specifically, "(1) to each person an equal share, (2) to each person according to individual need, (3) to each person according to individual effort, (4) to each person according to societal contribution, and (5) to each person according to merit" (p. 9). The authors also note that exposing a disadvantaged group to risk is an injustice (DHEW, 1978, pp. 6-7).

Fairness works as an ethical principle only if presented from within a certain set of parameters. These parameters were created out of thin air by Rawls (1971) through the device of a "veil of ignorance" (p. xix). They constitute the assumption that, given a choice, most people would choose a society much like the one we already have, something like a "meritocracy". That's an easy choice for a 1970s university-educated American. The reality is much different for people around the world.

"Fairness" in artificial intelligence amounts to the desire to, from a position of privilege, set those parameters that define where we will be "fair" and where we will make actual ethical decisions. Those with privilege will define "fairness" one way; those without privilege will define it very differently.

### 3.2 Transparency and explainability

It is argued that the use of AI is ethically problematic in society when it is not transparent. When a decision-making system is opaque, it is not possible to evaluate whether it is making the right decision. You might not even know the decision was made by a machine. Analytics requires a "principle of notification" (Fjeld et.al., 2020).

Additionally, transparency applies to the model or algorithm applied in analytics. "Transparency of models: it relates to the documentation of the AI processing chain, including the technical principles of the model, and the description of the data used for the conception of the model. This also encompasses elements that provide a good understanding of the model, and are related to the interpretability and explainability of models" (Hamon et al., 2020).

The intuition behind this condition is that AI systems should be designed in such a way that allows users to understand how they work. This can help users trust the systems they are using and to identify potential biases or errors. This, however, depends on the population as a whole having a reasonably sophisticated understanding of math and science, which may be too much to expect.

Even more to the point, this proposal amounts to the assertion that “AI systems should be designed differently than they are”. Most people understand the world in terms of simple principles and truisms, folk psychology and folk science (Natale & Ballatore, 2017). Causation is not complex in their world (there is no complexity in a world where we “understand” everything). However, as is well known, AI systems are exposed to datasets consisting of millions or even billions of parameters, and through multiple iterations of combining and filtering arrays of these values are able to detect regularities. Any human could understand how AI works because the process involves nothing more than basic math; it would, however, take several lifetimes to check the calculations at human speed.

The challenge of transparency and explainability is arguably as difficult for AI as it is for understanding how and why humans act the way they do, and as with humans, in practice actual transparency and explainability aren't that helpful. What we really want is some way to translate what went on in the other human's head to what goes on in ours, and maybe get a story in terms of beliefs, motives, intentions, and expectations (Thellman et al., 2017). This, though, is exactly what AI ethicists do not want when it comes to machines, so we are left in a bit of a quandary. We anthropomorphize even simple machines because it's how we understand complex phenomena.

In practice, our demands for transparency and explainability leave us with data and information that may or may not be relevant to our understanding. The requirement of transparency, for example, may tell us the source and nature of the training data, which we will then (by intuition) describe as “fair” or “representative”, without being in any sense able to comprehend the billions of bits of information. We say, for example, that the computer “reads” or “copies” the content, when it does no such thing (Vaswani et al., 2017). The meaning we see in a body of data is completely ignored by the computer.

Similarly, a prime candidate for “explainability” of AI processes is the postulation of counterfactuals (Chou, 2024). The computer produces a result “A” and we ask, what would have produced the result “B”? The semantics of counterfactuals are notoriously murky, however, involving at a minimum a predefinition of context and range of alternative possibilities, and at maximum a range of relevant possible worlds (Lewis, 1973). What counts as “most salient” and “most similar” in the analysis of a counterfactual statement very much depends on one's point of view.

Finally, it is arguable that our being able to see or explain the actions of an AI has nothing to do with the morality of the AI or its use. Most people neither see nor can explain internal combustion, and being able to do so has little or no bearing on their statements on the ethics of driving cars and trucks. Similarly, it literally doesn't matter how the sausage is made for people to decide whether or not they enjoy a good bratwurst with their beer.

### 3.3 Accountability

One of the motivations of principles like transparency and explainability is that there should be clear lines of responsibility for the development, deployment, and use of AI systems. The idea is that accountability would help ensure that AI is used responsibly and that those responsible can be held liable for any negative consequences. A number of agencies have announced efforts to ensure that automated decisions are “accountable” (Rieke et al., 2018). For example, the United Kingdom Data Ethics Framework argues for “setting the highest standards for transparency and accountability when building or buying new data technology” (Gov.UK, 2018).

A cynic might say that this amounts to the claim that there are no ethics without enforcement. It is the suggestion that unless people are held in some way responsible for what they will do with

AI, they will do things that are ethically wrong with AI (in this case, where “wrong” is interpreted as having “negative consequences”). The relation between ethics and enforcement is complex. While some may say “AI ethical principles have failed due to the lack of consequences” (Munn, 2022, p. 871) others may say that ethics that depends on enforcement is no ethics at all; as Kant (1785) would say, moral action arises from duty, for example, not from fear of punishment. It is the domain of law.

In any case, it's not clear that people actually value accountability. If they did, there would be no case for privacy and secrecy, because people would need to be accountable for what they did, whether it was with their money or in the bedroom. When people demand “accountability”, a cynic might say, it's usually only for certain things, and usually for things other people are doing.

But in fact, people often are held accountable in society (at least, poor people and scapegoats are held accountable; the more privileged you are, the less accountable you are). Commit a crime, cheat on a test, and you'll be held accountable (in theory) whether or not you use AI. So what is special about accountability where AI is concerned? How is it different to “use AI responsibly” than to “use a hammer responsibly” or “use a gun responsibly”?

The suggestion here is that there is some set of unethical things that can only be done with AI, or conversely, some things are unethical only if done with AI. However, these (ethically) amount to the same thing because there's nothing AI can do that a human can't do, given enough time and resources. Still, it is arguable that there is a difference. So what sort of things would make such a difference?

Mostly, they're the sorts of things really fast processing enables. For example, nobody has a problem if somebody counts every instance of “and then” in *War and Peace*, but if an AI does the same thing, it's represented as a form of copyright violation. Similarly, if a person watches everything you do in a public square, and recognizes you whenever you visit, that's fine, but if a computer does it, it's surveillance. Or if a person collects all public information about you and compiles a profile, that's called “good police work” or perhaps “crediting monitoring”, but if a computer does it, it's surveillance.

The difference isn't that humans don't do these things. What AI does, arguably, is that it amplifies these effects. But this is only a matter of scale, and the only difference is that humans, rather than AI, are employed to produce these effects, then they have to be paid, so only rich people or corporations can do them. If they are enabled by AI, the cost is greatly diminished, and anyone can do them. And that seems to be the point where it becomes a problem. But what sort of ethics is it where it's right if you are able to pay humans to do it, but wrong if you can program a computer to do it?

Finally, it is arguable that the notion of accountability is a fiction in any case. Most of the time we don't really want to know who is accountable, because that typically involves too many people. This is especially the case for widespread social issues like racism and climate change. Who, really, should be accountable for these? The notion of accountability identifies a problem and then picks out a specific wrong that is the cause of that problem. But the specific wrong is often the wrong that is easiest to identify or the most efficiently prosecuted. That's why even though we would like to hold billionaires accountable for their crimes, governments seem much more interested in the accountability of single parents on welfare (Imran, 2022; Bonn, 2017).

### 3.4 Privacy and data protection

It is often argued that AI systems should protect individuals' personal data and respect their privacy, and as the summary above notes, this includes ensuring that data is collected and used ethically and that individuals have control over their data. This connection is often made explicit;

Metcalf (2014), for example, classifies privacy under the heading of “respect for persons” (p. 2) while Baker et al. (2016) argue “Individuals’ rights include privacy, autonomy, and the ability to choose for themselves how they want to manage risk, consistent with their own personal values and life situations.”

It should be noted at the outset that this principle directly contradicts the one just above. We can have accountability, or we can have data protection, but we can’t have both at the same time for the same data. If privacy is unambiguously good, as is often argued (e.g., Dufresne, 2023), then individuals cannot be held accountable for their actions. There are many cases where “respect for persons” or “individual freedoms” will be abridged in the service of wider social goods; we are typically required to disclose our finances, our criminal past, or our manufacturing processes, as the case may be.

I’ve encountered this in discussions about research ethics. A researcher collects data about, say, an indigenous community in the Canadian far north. The researcher makes promises to the effect that the data will be destroyed after the research is completed, thus respecting the subject’s wishes. When the time to destroy the data comes, though, another department intervenes, and says the data must be retained. Otherwise, it would not be possible to hold researchers accountable if there were any human rights violations.

So, which right prevails? Arguably, this is usually resolved by determining who has more power, and not by determining who is more ethically right. Minimally, it could be argued that this is a political decision, based on which actor’s needs are most pressing at the time.

A second point could be raised here as well. What it is about AI that is special when it comes to privacy and data protection. People raised issues around privacy and data protection long before AI came on to the scene. How does AI make them different? After all, AI can’t do anything a human can’t do; it just does it a lot faster, and with far more data, and so issues around privacy would seem to be issues of scale and broader access, not something new in kind.

The response is that AI can draw inferences that are a lot harder (and more expensive) for humans to draw, as previously discussed. An AI, for example, can recognize a person by their gait. Humans can do this too, but only for people they already know well; an AI can do it for anyone. But again, though, we are forced to acknowledge that AI simply makes it possible for everybody to do something that was previously only the domain of rich people and corporations. It was OK when it was just the rich, but it’s not OK for everybody.

That’s why, for example, it’s OK for Equifax to connect data from a thousand sources to draw a complete financial profile of you and use that profile to tell people to deny you credit (DeMatteo, 2024); but it’s not OK for you as an individual to do the same thing with an AI and use that profile to deny them a date. It’s a funny idea of an ethical principle.

Finally, and this is similar to some of the objections raised above, the idea of protecting personal data has its downsides. Should a person be able to protect their criminal record from being observed? Was it wrong when the Panama Papers revealed tax evasion and corruption on the part of officials around the world? Is it ethical for an infected person to keep the fact of their contagious illness a secret? Should people have the right to hide their license plate information?

In fact, society is composed of a complex interplay of information we keep secret, information we share with only a few others, and information that the public has a right to know. Often data is co-created and hence subject to an even more complex set of principles. There are no absolutes here, and the boundaries are subject to negotiation and change. Different societies with different values draw the lines differently. The real ethical decisions are made when we make these decisions.

### 3.5 Inclusiveness

It is often argued that AI systems should be designed to be accessible to all users, regardless of their background or abilities. Issues of accessibility were raised even prior to the current wave of generative AI (e.g., Goldenthal et al., 2021). Accessibility has many dimensions, including issues related to disabilities, cost, technology, education and literacy, culture, and language.

It is also not reliably recognized as an ethical principle. Globally, society cannot even agree that everyone should have access to food, drinking water, and shelter. Almost nothing we produce is required to be used to benefit all members of society. It's hard to see how inclusiveness stands as an ethical value specifically for AI.

Indeed, it is arguable that the actual ethic in play in society is an ethic of unequal distribution of wealth and resources (including AI). The idea – and this is expressed more times than I can count – is that without unequal distribution, there is no incentive for people to make their lives, or society, better. “People should receive compensation congruent with their contributions,” writes Mankiw (2013, p. 32), among many others.

Even in wealthy societies, people who are deemed able to work but who are not working (for whatever reason) are allotted only the minimal allocation of welfare, an amount that by any standard has to be considered as punitive, on the grounds that they could be doing more, but they aren't. It's hard to imagine a change in the ethos of our society that would allocate the benefits of AI to such people.

In other countries, where there is not enough wealth to support welfare, people simply go hungry and homeless. Here, the welfare standard is applied on a national scale, where a certain amount of minimal foreign aid (usually provided on terms beneficial to the donor) is provided, a punitive amount, really, on the presumption that otherwise these nations (and the people in them) would have no incentive to improve their lot. Again, it's hard to imagine an “AI for Africa” campaign succeeding where “food for Africa” programs have failed.

Above, I mentioned the ethos of “taking care of one's own” (Hennessey, 2025). This applies here as well. The idea of inclusiveness suggests that we are all members of one society, but there are many people who believe that it is better to be members of exclusive societies. Certainly, this is an ethic of the wealthy, for whom exclusivity is a status symbol (Eastman et al., 2022).

There is also a sense of “algorithmic inclusivity” reflecting a concern that “discrimination and bias are inherent problems of many AI applications” (Fosch-Villaronga & Poulsen, 2022, p. 109). Groups facing discrimination either through unrepresentative data or structural bias may include women, racial minorities, the LGBTQ+ community, senior citizens, and disabled persons, among others. This may be, but it is far from universally held that such groups ought not face discrimination; for example, genders are not regarded equally in nations such as Iran and Afghanistan, and the principle of diversity, equity, and inclusion (DEI) has been explicitly rejected by the United States government.

If we, as a global society, really valued inclusion, we would be living in a very different world. But the history of the world is the history of the exertion of privilege by one group over another. I wish it were not the case, but I am not so I as to believe the rest of the world wishes it along with me.

### 3.6 Reliability and safety

Like everything else, AI systems should be reliable and safe to use. As software, two major risks are errors and malicious attacks. There are also warnings about the potential side-effects of AI in particular, ranging from lowered cognitive capacity to malevolent super-AI systems.

This is fair enough, but people are subject to much greater risks from other widely accepted technologies. For example, the design of cities, suburbs and workplaces requires that many employees commute to work, even in jobs they could perform equally well in the home. Return-to-office was widely mandated after the end of the Covid pandemic (Papamarko, 2025), even though driving kills roughly 2,000 people in Canada every year (Transport Canada, 2025). It's hard to reconcile this with a concept of "AI safety" from the same employer.

It's true that on balance we don't want to develop systems that kill people (unless, of course, we are developing systems to kill people, so they can be used by police or in war). We don't generally want to cause harm (unless, of course, we are in competition with someone, in which case causing harm may be fair). By the same token, we want the things we develop to be reliable, which means that they will do what we expect them to do, and not what we don't expect them to do, unless doing what we don't expect them to do is safer (ASQ, 2025).

As the exceptions and limitations of the previous paragraph demonstrate, the principles of reliability and safety are subject to a wide range of conditions, ranging from intent to expectations to trade-offs, cost, and efficiency. The principle, at best, can be expressed as "all else being equal, we want AI to be reliable, safe and secure."

If any discipline can bear witness to such conditions and trade-offs, it's computer science. Computers are reliable, but not in the way aircraft are reliable. Our computers are safe, but not in the way a bank vault is safe. On any given day, people wrestle with unintended consequences of their computers being neither reliable nor secure, often in ways that makes them unsafe – not in the sense of "you're gonna die" unsafe, but in the sense of "you're gonna lose your life savings" unsafe.

There is abundant evidence that we as a society are willing to be more than flexible when it comes to reliability and safety. Indeed, trust is more important than risk in shaping attitudes (Mou et al., 2017). More often than not, there is often widespread opposition to measures that make things more reliable and safe. In my lifetime, I've seen anti-seat-belt campaigns, anti-vaccination campaigns, a pro-gun lobby, and... well, you name it. It's not that being unsafe is an ethos (though for a certain segment of the population, it's certainly an ethos), it's that being safe is not the universal value people suppose it is (Hansson, 2023).

The same applies to resistance to malicious attacks. If this were really a universal value, nobody would ever set their password to 'password123' or write it down on a sticky note. Indeed, nobody would use the web at all. People gladly exchange safety and security all the time. It is inevitable that they will make the same trade-off when it comes to AI. And again, that's where the actual decisions about ethics will be made.

The argument, then, is not that AI and related technologies are without risks (though intuitively, these risks are less than those of technologies such as internal combustion engines and nuclear weapons). The argument is that, contrary to those who argue for first considering the harms and issues (Monett & Paquet, 2025), society as a rule tends to balance considerable levels of risk against anticipated benefits and trust in the provider. The suggestion that we should change this approach now, for this technology, does not appear to be any longstanding ethical principle but rather a latent conservatism that mistrusts new technologies in general, in other words, neo-Luddites (Lamont, 2024).

### 3.7 Human-centered design

It is often argued that AI systems should be designed specifically with the needs of humans in mind. This includes ensuring that AI systems are easy to use, understandable, and beneficial.

Human-centered design is not at first glance an ethical principle, except perhaps as opposed to the philosophies of people like John Muir (1897) and Katarzyna de Lazari-Radek and Peter Singer (2014, p. 380) to the effect that not everything needs to be designed with the needs of humans in mind. We also need to design and develop systems with other species, the environment, and possibly even the needs of artificial intelligences in mind (Gibert & Martin, 2022).

But the main point being made here is analogous to the thinking of ‘human-in-the-loop’ style arguments. The idea is that, ultimately, we want the humans to be in control, and the machines serving our needs. The human-in-the-loop argument could be understood in several different ways (Chiodo et al., 2025), but ultimately argues humans are necessary to ensure accountability and legal responsibility.

There are a couple of problems with this, though.

First, it runs against empirical evidence. It has often been observed that “the human shapes the tool; the tool then shapes the human” (Culkin, 1967, pp. 51-53). Whatever we build, we will adapt to it. There’s no such thing as a design that is purely “human-centered”. It’s an iterative process, where the imperatives and the affordances of human and machine shape each other. This has nothing to do with ethics; it’s just a fact about how interactive systems work.

Second, it’s a bad idea. There are many cases where it makes sense to have the machine over-rule the human, if at least only temporarily. My car, for example, won’t let me back into an obstacle. Instead, it triggers an emergency brake, which stops the car before it hits anything. This has no doubt saved the lives of countless children from the carelessness of their parents (Lorio, 2022).

Many things are like that. The needs of humans are overruled by machines all the time. The ATM won’t let me spend more money than I have. The gas pump will not let me have gas for free, even if I really need it. Even the most inert of technologies prevent me from jumping off high bridges or walking in through an outdoor. Systems designed like this reverse the logic of human-in-the-loop: as AI-in-the-loop, they offer valuable support and corrections for our everyday activities (Natarajan et al., 2025).

The same sort of case can be made for AI. Many baseball fans would like the final arbiter of balls and strikes to be a machine. The reason is that humans have historically done an awful job at this, even after years of training. Many would like AI systems to mark tests and evaluate job applicants because they are far less likely to favour their friends, discriminate against people with foreign names, or have a bad Monday (Chai et al., 2024).

Moreover, putting humans in control means keeping the humans who are in control, in control. There are so many cases where that’s a bad idea, I cannot even begin to list them all. Robbins (2024) warns, “In the race to replace ‘dumb’ infrastructure with ‘smart’ infrastructure, it is important that the new infrastructure is under society’s control” (p. 1383). What counts as, or constitutes, “society’s control”, is an open – and very political – question.

### 3.8 Non-maleficence

This principle argues that AI systems should be designed to avoid harming individuals, society, or the environment. This includes considering the potential risks and consequences of AI systems before they are deployed.

To point out the obvious: we very often design things specifically designed to harm individuals, society and the environment. They’re currently being deployed in places like Gaza and Ukraine. Major industries and billions of dollars are devoted to them. So, it matters very much what we decide is worth harming and what we decide should not be harmed.

I, personally, would have included babies and children in the “not to be harmed” category, but large numbers of people apparently disagree with me (Saugstad et al., 2024).

Having said that, as discussed previously, it is rare where we should not consider the potential risks and consequences. As a cyclist, I pay close attention to this. Similarly, when I’m engaged in my unnecessary commute, I’m careful to keep risks and consequences in mind. Most of the time, it’s pretty benign. We recently bought a new microwave; the biggest risks were that (a) it might not work, and (b) it might be from the U.S., which we’re boycotting. I also considered the possibility that it might explode.

It may be suggested that people generally say things like “This includes considering the potential risks and consequences” as an indirect way of saying “there are serious risks and consequences”. This may or may not be true as a matter of fact. But the ethical principle here is that it is best to avoid harm, or even more directly, that we should not deliberately cause harm. Unless, of course, we are seeking to deliberately cause harm.

The principle of non-maleficence originates in health care and generally reflects the idea that the treatment should not be worse than the disease. It is often carefully worded, because often one person’s medical procedure is another person’s stabbing. Context and outcome really matter. There are necessary harms, allowable harms, accidental harms, and unintended harms. All of these come into play, and how we determine whether a harm is unethical often depends on our point of view. Ask any person who has been laid off their job.

It is arguable that this ethical principle amounts to “don’t be evil”, but as we have all learned recently, it may not actually be legal to embrace that principle if it runs counter to shareholder interests, at least in some jurisdictions (Crofts & van Rijswijk, 2024).

### 3.9 Sustainability

It is frequently cited as an ethical principle that AI systems should be designed to be sustainable and environmentally responsible, which includes considering the energy consumption, water consumption, and other environmental impacts of AI systems. Environmental sustainability is a relatively recent arrival in the history of ethics, arguably originating with the works of Macpherson (1773) and Thoreau (1854).

Today, sustainability broadly conceived refers to the Sustainable Development Goals (SDG) (United Nations, 2015), though most of us reduce “sustainability” to “environmental responsibility”. Though it is frequently cited as an ethical responsibility, the best evidence (Lai, 2021) suggests that it isn’t for most people (and certainly not one that overrides, say, personal self-interest).

While we may be sympathetic with the idea that AI should be environmentally responsible, the fact remains that it’s among the least of our environmental concerns. Many things have far more impact. If I could work from home, I would be using far less energy, even if I did nothing other than use AI. If I did without my morning coffee, I would save more power than by quitting AI (Downey, 2025). The power consumed by AI is not even a rounding error when compared to the total amount of energy consumed by humans. It’s far too small to be considered even that.

In addition, let’s compare the energy we’re consuming if we’re not using AI. How many translators are there in the world. A (moderately trustworthy) Google search tells me there are 600,000 translators. Think of how much energy and resources it takes to pay for and support 600,000 translators. Though not quite there yet, it appears that AI will replace most of those. That means that whatever energy AI is using is replacing the energy we are using to support 600,000 translators.

Now, true, we are not actually eliminating 600,000 people (that would be quite unethical). With any luck (and an actual system of ethical stewardship), these 600,000 people would live their lives in more enjoyable and gainful pursuits. But that is worth the expenditure of energy, isn't it? There's a certain selfishness in the whole energy argument. Things are going well for us, so we shouldn't be expending more energy. But raising another billion people out of poverty will increase our energy consumption, so really, we shouldn't do it. It's the ethical thing to (not) do

## 4 Concluding remarks

University professors and the institutions they represent are frequently among those who define and recommend adherence to ethical principles in all we do, including the deployment of AI. That is laudable, but it's hard not to observe a certain amount of misdirected energy on the part of academia.

No, ethics does not consist of any of the principles listed above, nor of the several others that could be added from the many ethical treatises and frameworks addressing the use of artificial intelligence at educational institutions that have proliferated over the last few years. Many educators are very concerned that individuals, organizations, and institutions might use artificial intelligence to address the many problems people face in the world today.

That does not mean the proponents are wrong to take ethical factors into consideration. The selection of ethical arguments to apply, however, and the expression of ethical principles in the "consensus list" represents just one of many possible points of view. Many "ethical" principles are also social, cultural and political in nature, and there is a wide variance of opinion on what's important and what is ethically relevant.

For example, we live in a world of poverty and war, famine and hardship. We live in a world where there is injustice and deprivation, where corruption and oppression are so often the order of the day. And amid that, the very same people loudly proclaim that they have solved ethics, and that we should follow their lead.

And that's fine, so far as it goes, until what they decide is ethical is a standard that ought to apply to everyone. And then, from my perspective, it's like faith standing in for science, and we get pronouncements like "there must be an explanation" or "somebody must be responsible". And that's when it becomes necessary to push back.

Without raising any false hopes and without excessive hype, it is already clear that the use of artificial intelligence, including generative AI, could open access and improve learning for billions of people who are otherwise unable to benefit from the benefits that a modern information-age civilization can provide. It is difficult to countenance arguments against the use of AI on political grounds, much less ethical grounds, in the face of this reality.

Concerns about the use of AI are fairly raised as political arguments, suggesting that the direction and funding of great institutions ought to flow in one direction or another. But they are not ethical arguments. Ethics isn't at all about what others should do. That's a matter for jurisprudence and law, the systems we have struggled to build over the years to allow us to live alongside each other without (too much) fear and violence.

And indeed, most of the treatises and frameworks that pose as ethical principles are actually about laws and regulations. What kind of behaviour can we convince other people to follow? Or if we can't convince them, what can we put into law (with a suitable accountability framework) so we can force them to follow? That, though, has nothing to do with ethics. It's politics.

I have nothing against politics, except when it pretends to be ethics. Politics isn't about what's right, because we don't agree on what's right. That's why we need politics.

Ethics, by contrast, is personal. It's based in our own sense of what's right and what's wrong (itself a product of culture and education and upbringing and experience and reflection) and is manifest in different ways in different people (and not at all in psychopaths) and for me is a combination of empathy and fear and loathing and – on my good days – of peace and harmony and balance. It consists of what I am willing to allow of myself, what guides my decisions, what I am willing to accept, and what will cause me to push back with a little force or all the might I possess.

The argument that claims there is “consensus” around the ethical application of AI in education, or any other domain, fails to take these considerations into account. It substitutes a specific ethical perspective to stand in the place of political debate, and as a consequence sidelines some of the very real issues that apply to each of the principles listed and discussed in this paper.

## References

- ASQ (American Society for Quality). (2025). What is reliability? <https://asq.org/quality-resources/reliability>
- Baker, D. B., Kaye, J., & Terry, S. F. (2016). Governance through privacy, fairness, and respect for individuals. *EGEMS (Wash DC)*, 4(2), 1207. <https://doi.org/10.13063/2327-9214.1207>
- Baker, S., & Inventado, P. S. (2016). Educational data mining and learning analytics: Potentials and possibilities for online education. In G. Veletsianos (Ed.), *Emergence and innovation in digital learning* (pp. 83–98). Athabasca University Press. <https://doi.org/10.15215/aupress/9781771991490.01>
- Bonn, S.A. (2017, April). Why elite white-collar criminals are rarely punished. *Psychology Today*. <https://www.psychologytoday.com/us/blog/wicked-deeds/201704/why-elite-white-collar-criminals-are-rarely-punished>
- Cardona, M., Rodríguez, R.J., & Ishmael, K. (2023). *Artificial intelligence and future of teaching and learning: Insights and recommendations*. U.S. Department of Education, Office of Educational Technology. <https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf>
- Chai, F., Ma, J., Wang, Y., Zhu, J., & Han, T. (2024). Grading by AI makes me feel fairer? How different evaluators affect college students' perception of fairness. *Frontiers in Psychology*, 15, 1221177. <https://doi.org/10.3389/fpsyg.2024.1221177>
- Chiodo, M., Muller, D., Siewert, P., Wetherall, J., Yasmine, Z., & Burden, J. (2025). Formalising human-in-the-loop: computational reductions, failure modes, and legal-moral responsibility. <https://arxiv.org/pdf/2505.10426>
- Chou, Y. (2024). *Counterfactuals in explainable artificial intelligence: Theory, algorithms, and applications [Unpublished Master's thesis]*. School of Information Systems, Faculty of Science, Queensland University of Technology. [https://eprints.qut.edu.au/248281/1/Yu-Liang+\(Leon\)+Chou+Thesis.pdf](https://eprints.qut.edu.au/248281/1/Yu-Liang+(Leon)+Chou+Thesis.pdf)
- Culkin, J. M. (1967, March 18). A schoolman's guide to Marshall McLuhan. *Saturday Review*. <https://webpace.royalroads.ca/llefevre/wp-content/uploads/sites/258/2017/08/A-Schoolmans-Guide-to-Marshall-McLuhan-1.pdf>

- Crofts, P., & van Rijswijk, H. (2020). Negotiating 'evil': Google, Project Maven and the corporate form. *Law, Technology and Humans*, 2(1), 75-90. <https://doi.org/10.5204/lthj.v2i1.1313>
- de Lazari-Radek, K., & Singer, P. (2014). *The point of view of the universe: Sidgwick and contemporary ethics*. Oxford University Press. <https://www.utilitarianism.com/view-universe.pdf>
- DeMatteo, M. (2024, Aug 7). The credit score mortgage lenders look at when you apply for a home loan. *CNBC*. <https://www.cnbc.com/select/which-credit-score-used-when-applying-for-mortgage/>
- DHEW (Department of Health, Education and Welfare). (1978). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research, Report of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research*. [http://videocast.nih.gov/pdf/ohrp\\_belmont\\_report.pdf](http://videocast.nih.gov/pdf/ohrp_belmont_report.pdf)
- Downey, S. (2025). AI, energy use, and the coffee dilemma. *LinkedIn*. <https://www.linkedin.com/pulse/ai-energy-use-coffee-dilemma-sarah-downey-i48pc/>
- Drenon, B. (2025, March 25). In the US, DEI is under attack. But under a different name, it might live on. *BBC*. <https://www.bbc.com/news/articles/c24110m30ddo>
- Dufresne, P. (2023). Privacy as a fundamental right in the digital age. Keynote remarks at the 25<sup>th</sup> Annual Vancouver International Privacy & Security Summit (VIPSS). February 24, 2023, Vancouver, British Columbia. Office of the Privacy Commissioner of Canada. [https://www.priv.gc.ca/en/opc-news/speeches-and-statements/2023/sp-d\\_20230224/](https://www.priv.gc.ca/en/opc-news/speeches-and-statements/2023/sp-d_20230224/)
- Eastman, J. K., Iyer, R., & Babin, B. (2022). Luxury not for the masses: Measuring inconspicuous luxury motivations. *Journal of Business Research*, 145, 509-523. <https://doi.org/10.1016/j.jbusres.2022.03.022>
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI* (Berkman Klein Center for Internet & Society). [https://dash.harvard.edu/bitstream/handle/1/42160420/HLS%20White%20Paper%20Final\\_v3.pdf?sequence=1&isAllowed=y](https://dash.harvard.edu/bitstream/handle/1/42160420/HLS%20White%20Paper%20Final_v3.pdf?sequence=1&isAllowed=y)
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People – An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Fosch-Villaronga, E., & Poulsen, A. (2022). Diversity and inclusion in artificial intelligence. In B. Custers & E. Fosch-Villaronga (Eds.), *Law and artificial intelligence*. Information Technology and Law Series, vol 35, pp. 109-134. T.M.C. Asser Press. [https://doi.org/10.1007/978-94-6265-523-2\\_6](https://doi.org/10.1007/978-94-6265-523-2_6)
- Gibert, M., Martin, D. (2022). In search of the moral status of AI: Why sentience is a strong argument. *AI & Society*, 37, 319–330. <https://doi.org/10.1007/s00146-021-01179-z>
- Goldenthal, E., Park, J., Liu, S. X., Mieczkowski, H., & Hancock, J. T. (2021). Not all AI are equal: Exploring the accessibility of AI-mediated communication technology. *Computers in Human Behavior*, 125, 106975. <https://doi.org/10.1016/j.chb.2021.106975>
- Gov.UK Open Research Data Task Force. (2018). *Realising the potential: Final report of the Open Research Data Task Force*. Jisc. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/775006/Realising-the-potential-ORDTF-July-2018.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/775006/Realising-the-potential-ORDTF-July-2018.pdf)
- Hakimi, L., Eynon, R., & Murphy, V. A. (2021). The ethics of using digital trace data in education: A thematic review of the research landscape. *Review of Educational Research*, 91(5), 671–717. <https://doi.org/10.3102%2F00346543211020116>

- Hamon, R., Junklewitz, H., & Sanchez, I. (2020). *Robustness and explainability of artificial intelligence—From technical to policy solutions* (No. 978-92-79-14660-5). Publications Office of the European Union. <https://doi.org/10.2760/57493>
- Hansson, S. O. (2023). Liberty, paternalism, and road safety. In K. Edvardsson Björnberg, S. O. Hansson, M. Å. Belin & C. Tingvall (Eds.), *The Vision Zero Handbook*. Springer, Cham. [https://doi.org/10.1007/978-3-030-23176-7\\_6-2](https://doi.org/10.1007/978-3-030-23176-7_6-2)
- Hennessey, M. (2025, June 30). Capitalism needs champions. *The Wall Street Journal*. <https://www.wsj.com/opinion/capitalism-needs-champions-free-markets-versus-socialism-5114b804?st=JRnSTu>
- Imran, A. (2022, April 28). A comparative analysis of sentencing disparities involving white-collar crime. *The Society: Sociology & Criminology Undergraduate Review*, 7(1). <https://jps.library.utoronto.ca/index.php/society/article/download/38489/29336/102293>
- Kant, I. (1785). *Groundwork of the metaphysics of morals*. Wikisource. [https://en.wikisource.org/wiki/Groundwork\\_of\\_the\\_Metaphysics\\_of\\_Morals](https://en.wikisource.org/wiki/Groundwork_of_the_Metaphysics_of_Morals)
- Kheya, T. A., Bouadjenek, M. R., & Aryal, S. (2024). The pursuit of fairness in artificial intelligence models: A survey. <https://arxiv.org/abs/2403.17333>
- Lai, O. (2021, September 17). Every G20 country is failing to meet Paris Agreement On Climate Change. *Earth.org*. <https://earth.org/every-g20-country-is-failing-to-meet-paris-agreement-on-climate-change/>
- Lamont, T. (2024, February 17). ‘Humanity’s remaining timeline? It looks more like five years than 50’: Meet the neo-luddites warning of an AI apocalypse. *The Guardian*. <https://www.theguardian.com/technology/2024/feb/17/humanitys-remaining-timeline-it-looks-more-like-five-years-than-50-meet-the-neo-luddites-warning-of-an-ai-apocalypse>
- Lewis, D. (1973). *Counterfactuals*. Blackwell. <https://perso.uclouvain.be/peter.verdee/counterfactuals/lewis.pdf>
- Lorio, J. (2022, November 16). Car safety tech really does cut number of rear-end crashes drastically. *Car and Driver*. <https://www.caranddriver.com/news/a41981175/automated-emergency-braking-forward-collision-warning-effectiveness/>
- Macpherson, J. (1773). *The poems of Ossian*. A new edition, carefully corrected, and greatly improved. <https://bellacaledonia.org.uk/2025/07/16/ossian-punk-beer-and-greenwash/>
- Mankiw, N. G. (2013). Defending the one percent. *Journal of Economic Perspectives*, 27(3), 21–34. <http://dx.doi.org/10.1257/jep.27.3.21>
- McRae, P., & Bower, J. (2013, April 30). Rebirth of the teaching machine through the seduction of data analytics: This time it's personal. *National Education Policy Center (NEPC)*. <https://nepc.colorado.edu/blog/rebirth-teaching-machine-through-seduction-data-analytics-time-its-personal>
- Metcalf, J. (2014). *Ethics codes: History, context, and challenges*. Council for Big Data, Ethics, and Society. <https://bdes.datasociety.net/wp-content/uploads/2016/10/EthicsCodes.pdf>
- Monett, D. & Paquet, G. (2025). Against the commodification of education — if harms then not AI. *Journal of Open, Distance, and Digital Education*, 2(1). <https://doi.org/10.25619/wazgw457>
- Mou, J., Shin, D. H., & Cohen, J. F. (2017). Trust and risk in consumer acceptance of e-services. *Electronic Commerce Research*, 17, 255–28. <https://doi.org/10.1007/s10660-015-9205-4>

- Muir, J. (1897). The American forests. *John Muir: A reading bibliography by Kimes*, 1986 (Muir articles 1866-1986). 234. <https://scholarlycommons.pacific.edu/jmb/234>
- Munn, L. (2023). The uselessness of AI ethics. *AI Ethics*, 3, 869–877. <https://doi.org/10.1007/s43681-022-00209-w>
- Natale, S., & Ballatore, A. (2017). Imagining the thinking machine: Technological myths and the rise of artificial intelligence. *Convergence*, 26(1), 3-18. <https://doi.org/10.1177/1354856517715164>
- Natarajan, S., Mathur, S., Sidheekh, S., Stammer, W., & Kersting, K. (2025). Human-in-the-loop or AI-in-the-loop? Automate or collaborate?. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27), 28594-28600. <https://doi.org/10.1609/aaai.v39i27.35083>
- Novak, M. (2022). Auto-Tutor (1964). *Paleofuture*. Retrieved 2025-07-02. <https://paleofuture.com/blog/2008/4/23/auto-tutor-1964.html>
- Papamarko, J. (2025, January 17). Stuck in the past? Why leaders are wrong about return-to-office mandates. *McMaster University*. <https://mcm.humanities.mcmaster.ca/2025/01/17/stuck-in-the-past-why-leaders-are-wrong-about-return-to-office-mandates/>
- Pemberton, N. (2022, August 29). Life isn't fair. Get over it. *Counterpunch*. <https://www.counterpunch.org/2022/08/29/life-isnt-fair-get-over-it/>
- Pressey, S. L. (1926). A simple apparatus which gives tests and scores—and teaches. *School & Society*, 23 (586), 373–376. ISSN 0036-6455. Reprinted in Lumsdaine & Glaser 1960, pp. 35–41.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press. Revised edition, 1999. <https://archive.org/details/theoryofjustice00rawl/page/n5/mode/2up>
- Regents of the University of Michigan (U of M). (2025). Learning analytics guiding principles. Center for Academic Innovation we page. <https://ai.umich.edu/learning-analytics-guiding-principles/>
- Rieke, A., Bogen, M., & Robinson, D. G. (2018). Public scrutiny of automated decisions: Early lessons and emerging methods. *Upturn and Omidyar Network*. <https://www.omidyar.com/insights/public-scrutiny-automated-decisions-early-lessons-and-emerging-methods>
- Robbins, S. (2024). The many meanings of meaningful human control. *AI Ethics*, 4, 1377–1388. <https://doi.org/10.1007/s43681-023-00320-6>
- Saugstad, O. D., Modi, N., Moretti, C., Obladen, M., Vento M., & Speer, C. P. (2024). Newborns and children in war and terror. *Neonatology*, 121(2), 137-140. 140. <https://doi.org/10.1159/000535401>
- Sghir, N., Adadi, A., & Lahmer, M. (2023). Recent advances in predictive learning analytics: A decade systematic review (2012–2022). *Education and Information Technologies*, 28, 8299–8333. <https://doi.org/10.1007/s10639-022-11536-0>
- Skinner, B. F. (1968). *The technology of teaching*. *The Century psychology series*. Appleton-Century-Crofts.
- SOLAR (Society for Learning Analytics Research). (2025). Definition of learning analytics. <https://www.solaresearch.org/about/what-is-learning-analytics/>
- Thellman, S., Silvervarg, A., & Ziemke, T. (2017). Folk-psychological interpretation of human vs. humanoid robot behavior: Exploring the intentional stance toward robots. *Frontiers in Psychology*, 8, 1962. <https://doi.org/10.3389/fpsyg.2017.01962>

- Thoreau, H. D. (1854). *Walden, or, life in the woods*. [https://en.wikisource.org/wiki/Walden\\_\(1854\)\\_Thoreau](https://en.wikisource.org/wiki/Walden_(1854)_Thoreau)
- Transport Canada. (2025). *Canadian motor vehicle traffic collision statistics: 2023*. Government of Canada. <https://tc.canada.ca/en/road-transportation/statistics-data/canadian-motor-vehicle-traffic-collision-statistics/2023/canadian-motor-vehicle-traffic-collision-statistics-2023>
- United Nations (UN). (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. <https://sdgs.un.org/2030agenda>
- University of Michigan. (2017). *Learning analytics guiding principles*. <https://ai.umich.edu/learning-analytics-guiding-principles/>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 6000-6010). [https://proceedings.neurips.cc/paper\\_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf)
- Ye, D. (2022). The history and development of learning analytics in learning, design, & technology field. *TechTrends*, 66, 607–615. <https://doi.org/10.1007/s11528-022-00720-1>

## Acknowledgement

Thanks to Junhong Xiao for recommending and motivating this publication.

## Author's Contributions (CRediT)

SD: conceptualization, investigation, writing - original draft, writing – review & editing, funding. The author has read and agreed to the published version of the manuscript.

## Competing Interests

The author has no competing interests to declare.

## Acknowledgement of Use of Generative AI Tools

GenAI was not used in the development of this article.