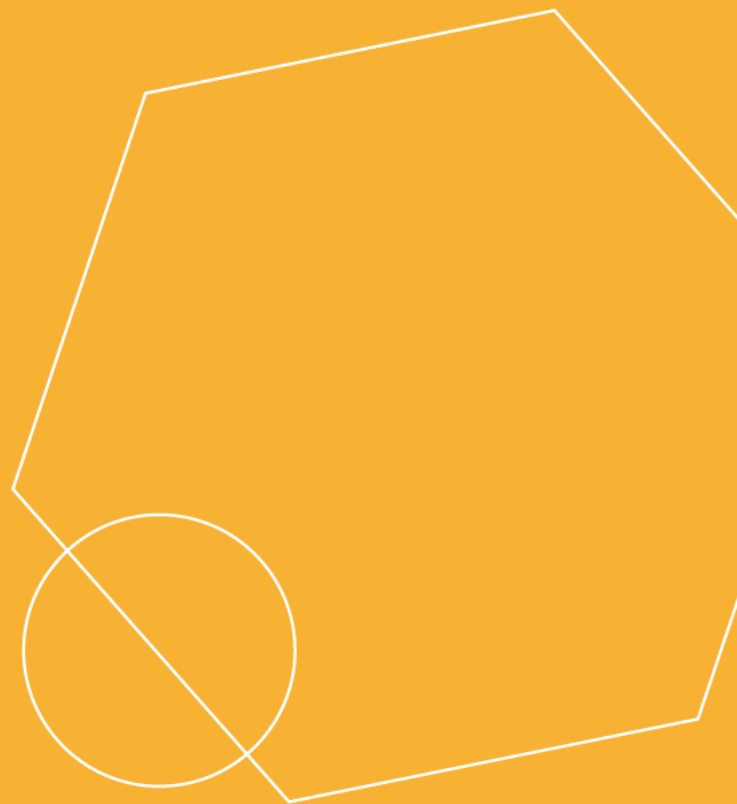


The AI Primer

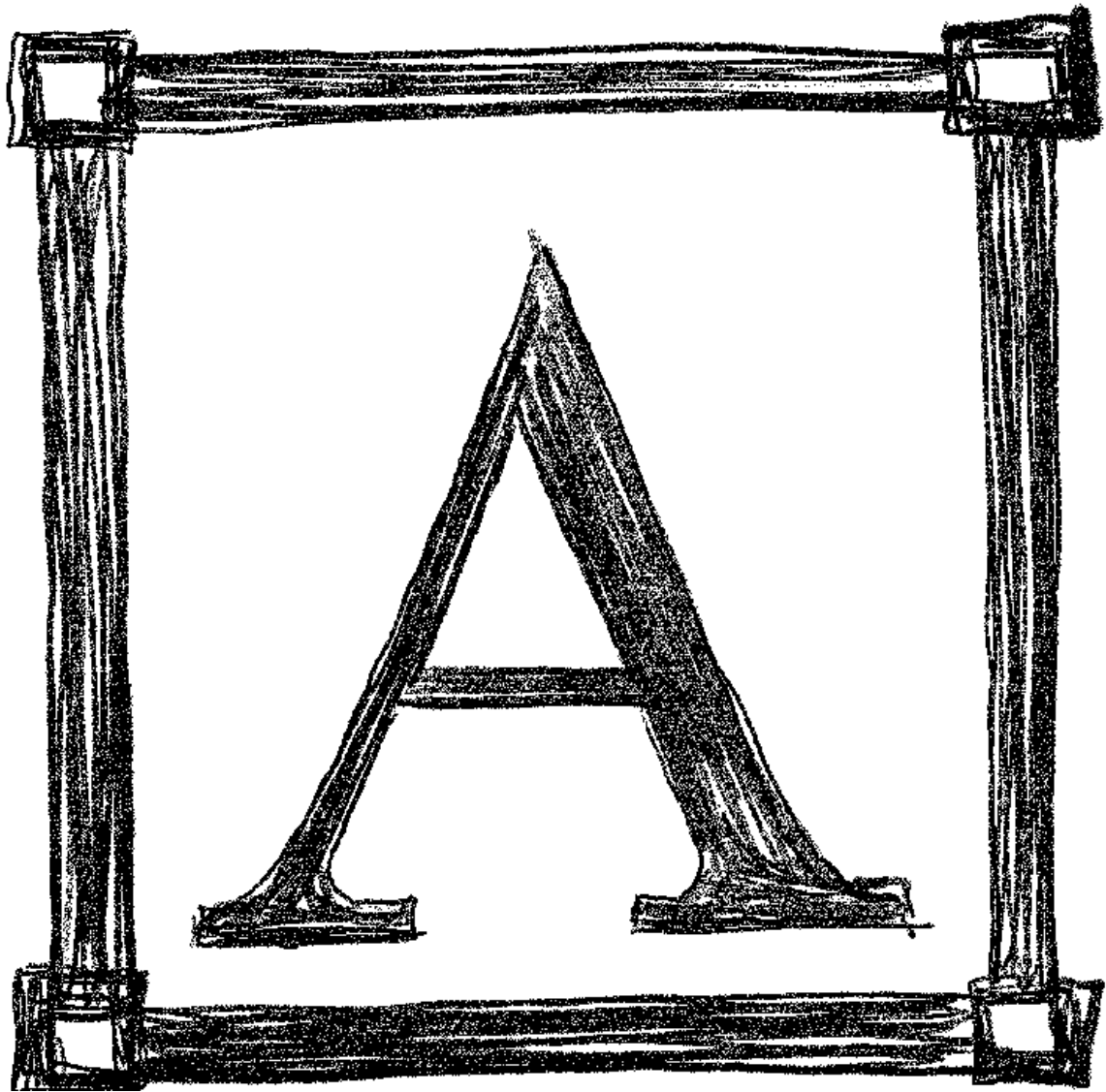
A Companion for developing AI literacy

By Juliet Waters

Illustrations by Danielle Grisnich



The ABC of citizenship in the era of AI



A is for attitude

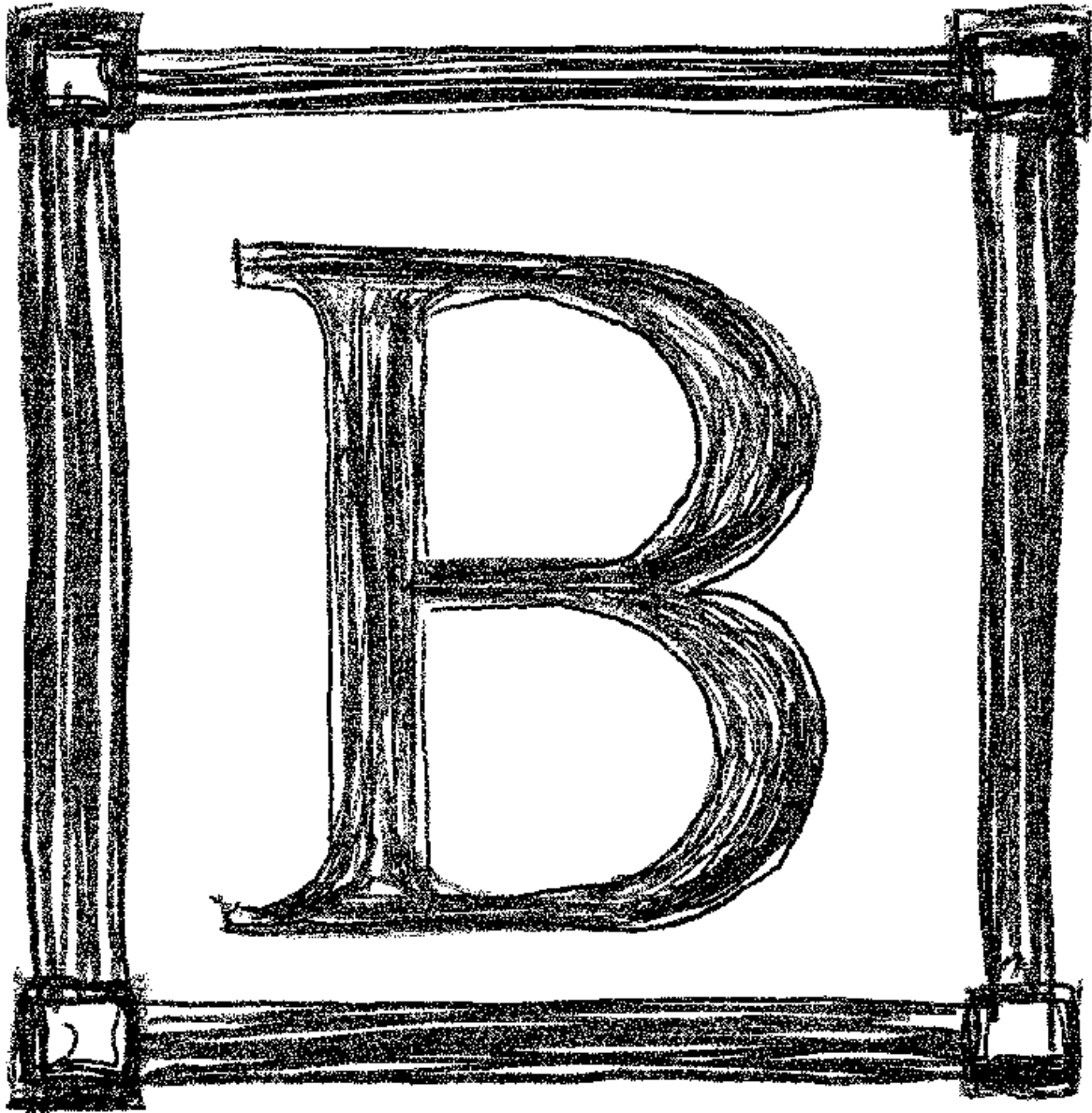
How does the idea of artificial intelligence make you feel? Are you curious or hesitant? Excited or worried? Do you look forward to the convenience of astonishing chatbots and self-driving cars? Do you picture yourself relegating mundane tasks to artificial intelligence (AI), and having more time and attention for family, friends, and stimulating activities? Or are you worried that AI might replace the skills and talents you bring to the world?

Rest assured that whatever you're feeling, you are not alone.

A recent survey of [120 countries](#) indicates that different regions feel differently about AI. In Europe, North America, and Latin America, negative attitudes dominate. In the Middle East as well as Central and South Asia, positive attitudes slightly outweigh negative ones. An even greater gap emerges in Africa and Southeast Asia, where views towards AI are more positive. In China, only 11% of respondents had negative attitudes towards AI.

The regional split in views about AI likely reflects cultural and economic differences across the globe, and this sparks a further question: do we form our opinions about AI based on knowledge and experience, or based on personal and cultural bias? Whatever our perspective, AI is not going away. Its impact on our society is accelerating. The changes it brings will affect the way that we interact with our economy as well as social relationships and citizen responsibilities. This occurs because all technological changes, including the invention of writing, the printing press, telephones, the internet, or AI, eventually impact the way we understand, use, and communicate information. Changes to information processing also change society. We owe it to ourselves to have an informed opinion about AI as well as a basic understanding of how it functions. This will help us work, teach, and live alongside intelligent machines.

Let's balance healthy skepticism and caution with determined open-mindedness and continual vigilance to re-evaluate our own preconceptions.



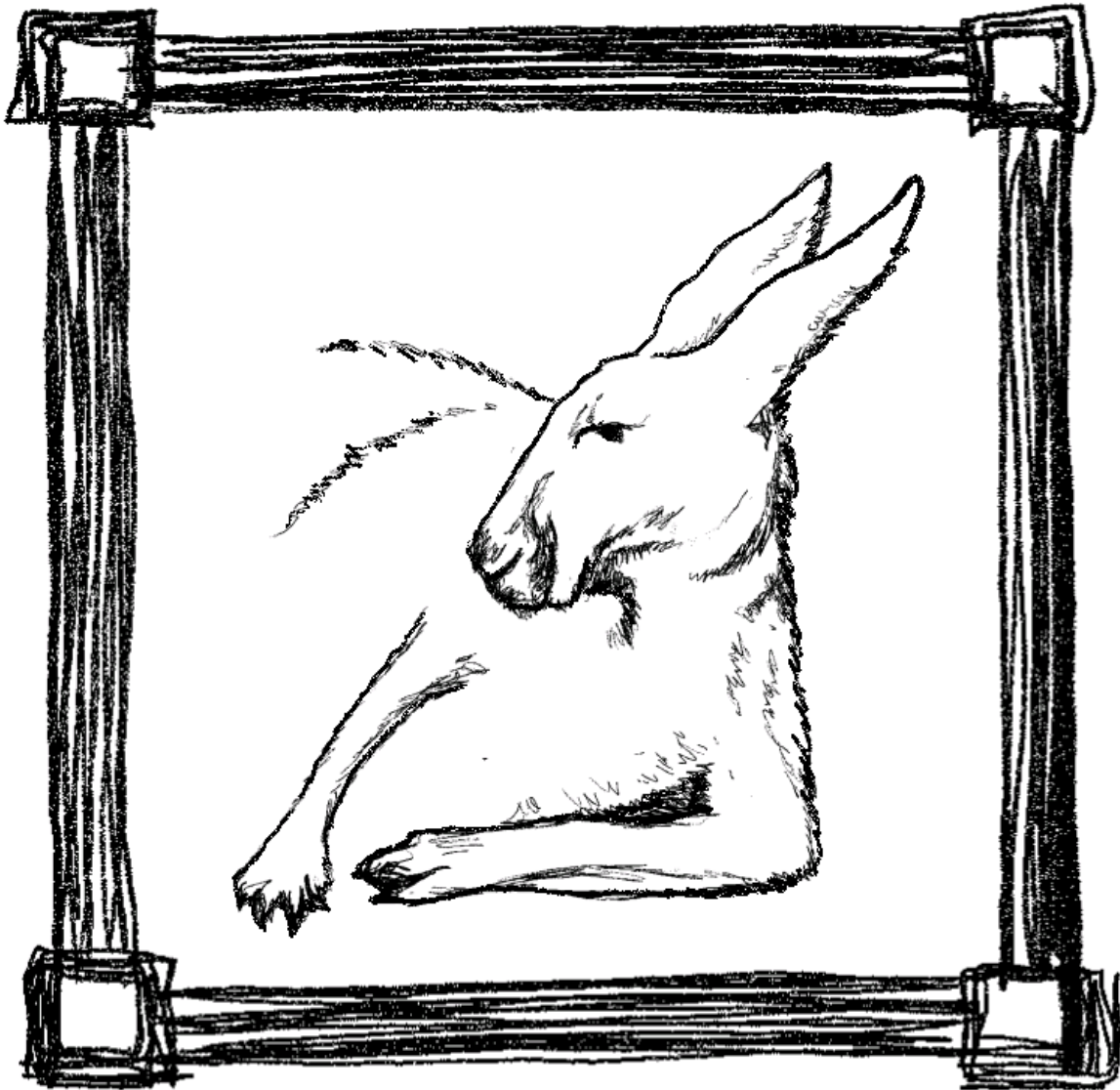
B is for bees, and our collective human brain power

In 2017, the United Nations declared May 20 World Bee Day. Thirty-five percent of bees die every year, largely due to accelerating climate change trends such as fluctuating seasonal patterns and more frequent natural disasters. Traditional beekeepers are simply not able to keep up with the massive amount of work it will take to restore that balance. This is very bad news, but not just for bees. With 90% of wild plants and 75% of crops dependent on pollination, one out of every three bitefuls of the food we eat requires their survival. Without bees, a third of the world food supply could disappear!

But here is some good news. Thanks to AI innovators, we now have solar-powered autonomous beehives with computer vision and robotics that help beekeepers monitor and thermoregulate their beehives, organically reduce pests, increase bee survival rates, and decrease the amount of human labour needed to care for bees.

If by 2030 bees are thriving, and food is growing at the rate we need it to keep growing, that may have a lot to do with the power of AI to accelerate positive change. But it will also demonstrate our collective human brain power as we develop ingenious solutions that harness AI to serve ecosystems and human well-being.

C is for caring about kangaroos and all species (including humans) impacted by AI



When Volvo began testing their [first self-driving cars in Australia](#) they ran into a problem. The vehicles could identify and avoid Northern European animals such as deer and elk. But to an AI that does not have intuitive depth perception, kangaroos

were a different problem. When kangaroos are in mid-flight they look like they are further away than they actually are. When they land, they seem closer. The AI had no way of explaining or correcting for this confusion. It was up to humans to figure it out and provide enough data on kangaroos to set the AI straight.

But not every challenge is as obvious to the human eye as the kangaroo identification problem. AI takes in a lot of data and recognizes complex patterns within that data, but it doesn't really understand the information and can't supply context.

For instance the AI assistant Alexa can learn to recognize a command like "order me a dollhouse." But it might not be able to distinguish between whether a parent or a child is ordering the dollhouse. This actually happened in one family, when it was discovered that their little girl had been able to order a very expensive dollhouse for herself. When the story was told one day on local radio by a commentator who said, "I love the story of the little girl who said Alexa, order a dollhouse," another Alexa in another house heard the command and ordered another dollhouse.

In the past computers followed rules created by human beings. But now, because of advances in "machine learning," AI is driven by data. To a large extent it makes up its own rules based on what it deciphers from this data. We will learn more about this process later.

AI's value lies in its ability to process enormous amounts of data. However, this very ability presents new problems. As the AI field progresses and learning algorithms become more complex, AI develops strategies that are impossible for a human, and even an AI system, to track.

This is called a "black box" problem.

You may have heard of the black box in airplanes or cars that are accessed after a crash. Those boxes record all the activities and commands leading up to an accident so that we can learn from what went wrong.

If only AI had a similar box! Unfortunately the amount of data that machine learning algorithms create, means that if there actually were an actual black box in an AI

system that we could open, it would probably just emit a lot of indecipherable white noise.

The black box problem in AI is that there is, effectively, no black box.

Machine Learning algorithms process so much information that AI is somewhat mysterious to us and even to itself. From where we sit, all we can really be sure of is what data goes into algorithms and what data comes out. This means that AI still requires human intervention to check and balance its work.

AI can do a lot to make the world safer and more efficient. It can also do a lot to make the world even more dangerous, confusing and inequitable than it already is.

Which leads us to one more letter...

D is for data, decisions, and diversity



In the end AI will only ever be as good, or as trustworthy, as the quality, the diversity and the explainability of the data it is trained on. Moreover, it will only be as valuable as our ability to make good decisions on how it should—or maybe should not—be used.

We all share the desire to be happy and to put our collective intelligence towards the task of building a world that nurtures safety, well being and mutual respect. How AI fits into that goal is up to us!

Getting Started

What is a primer?

First what it isn't: A primer is not a comprehensive guide. Nor is it a textbook that will tell you what's going to be on the exam. This is not where you will find a glossary of terms (though this primer will direct you to other resources that will include glossaries.)

A good primer is a friendly companion that makes those first steps easier. It builds on existing knowledge and provides just enough new knowledge to set you on a solid path to building more. Think of a classic literacy primer that uses a child's bank of pictures and oral memory and puts simple words to those pictures so that a child can start the process of decoding new knowledge.

But a primer for people who have already spent some time living and learning in the world has to do one other thing. It also has to identify existing knowledge and assumptions that might be conflicting with new ways of seeing things. Think of another meaning of primer: a coat of paint you have to put down so that a new colour of paint appears as you want it to and isn't distorted by the colours or patterns that were there before.

Especially in the world of AI, we have to be conscious of our assumptions. Our assumptions about society and reality creep into the way we make and use AI, and can become biases that diminish both the usefulness and the fairness of the AI.

MIT researcher Joy Buolamwini [discovered this](#) when the AI facial recognition algorithms she was working with failed to recognize her face. The algorithms hadn't been trained on enough pictures of Black women. The programmers who trained the AI assumed that the massive data banks of photos with which they had trained their AI reflected the diversity of the world. But they were wrong. No automated system is

inherently neutral. Every automated program, like all the things people make, reflects the knowledge and assumptions of the person who creates it. If the creators of an AI have subconscious false assumptions about people, the AI may entrench these errors in larger systems that impact thousands or even millions of people.

Citizenship in the era of AI starts with a mindset of intentionally confronting and mitigating our biases, especially when the systems around us might amplify them.

Who is this Primer for?

This primer is written for teens, educators, and adults who are curious about AI, in both formal and informal settings. It will help them:

1. Recognize what they already know.
2. Build a good conceptual understanding of what they have yet to learn, and perhaps unlearn.
3. Try some things out so that they can see how AI works first hand.
4. Feel empowered to share what they've learned with others.
5. Have fun!

Once this primer has helped you conceptualize how AI works, it will direct you to fun hands-on activities and further pathways to pursue, not just for learning, but for play.

AI can enrich our lives in all kinds of ways beyond just solving problems. AI is being used all over the world to create art, media, and countless other inventive, collaborative projects. AI can make life more colourful, inspiring, and convenient. It can give us more time to be great friends, parents, children and citizens. AI has the potential to be a tool of co-creation. But that outcome depends on balancing confidence and collaboration with respect for other people's perspectives, privacy and comfort level with technology.

Before we get started: Don't forget that black box!

No matter how much fun or success you're having with AI, never forget those kangaroos! Use them to keep in mind that no matter how much we learn, one problem remains, AI can never entirely explain how it solves problems. Its data sets are too big and its algorithms are too complex, which is why it's so important that humans retain oversight as we transition to a society more and more dependent on intelligent machines.



AI can provide answers, but unlike you, AI can't understand the answers.

Let's have AI explain this to you:



Does ChatGPT understand questions?





Yes, ChatGPT is able to understand and respond to questions. It is a language model that has been trained on a large dataset of text and can generate human-like responses to a wide variety of prompts. If you have any questions, feel free to ask.  



Does ChatGPT understand what it's writing?



ChatGPT, like other language models, does not have the ability to understand the meaning of the text it generates. It is trained to predict the next word in a sentence based on the patterns it has learned from the text data it was trained on. However, it can still generate coherent and often human-like responses to prompts, but it doesn't have any understanding of what it's writing.  

But how can I learn about AI if I may never be able to trace the process AI uses to come to its answers?

Again, it comes down to attitude. Just because you can't know everything doesn't mean you can't start and keep learning. You may not understand exactly how plants

grow, but that need not stop you from gardening. You may not fully understand how the brain learns, but that will not prevent you from becoming a teacher.

As you learn about and work with AI you will develop a conceptual understanding of how AI works. You will discover the best kinds of questions to ask of the AI you use. Learning to ask the right questions is essential for all citizens, whether students, non-experts, and even experts, because this helps us grasp the scope of the problems AI can solve. It helps us assess what risks AI poses. Knowing that, we can decide what problems might still be better solved using natural human intelligence, critical thought, interpersonal dialogue, and other traditional problem-solving methods.

We learn about AI so that we can make the best decisions about which activities, in which conditions, are appropriate places to use it. And we may well find that some activities and some conditions are not a good fit. If AI makes mistakes in recommending what movie it thinks we might like, it's unlikely we would decide not to use it, as long as it is usually reliable. If an AI elder care robot designed to transfer 90 year old patients from a wheelchair to a bed, only drops one out of every five patients, or even one out of every 1000—that is not AI we could ethically use without continuing to ensure that there was significant and ongoing human regulation and oversight.

So let's get started with the first concept we need to understand when talking about AI in the world we live in: Machine Learning.

The secret to learning machine learning

In one very important way, you probably already know how machine learning works!

If you've ever developed a preference for a particular kind of food, you probably developed it in much the same way that machines develop their learning skills. The big difference is that machines do not experience the sensations by which you build your food preferences. Machines can learn from images, sounds, numbers, and many other kinds of data, but that is not what we understand as "sensation."

For now, let's look at what we have in common with computers: a binary "source language" that helps us interpret the world. **Binary** refers to any scheme that is built on only two values. For computers, this binary source language is the pattern of 1s and 0s from which all other computer languages and code are built. Humans have something very similar to a binary source language in the way the brain categorizes sensations as pleasant or unpleasant. Simply from liking and not liking sensations, and tracking the degree to which we like or don't like sensations, our brains have created millions of responses to our world. Liking and disliking helps us navigate choice in our complex environments.

At a fundamental level this is how machine learning works. It's pretty simple. A computer takes in sets of data, labels the data, and then tries to create rules from the patterns it discovers.

Making a model

Another process we have in common with machines is that our learning creates models, and then we use those models to predict things.

Imagine a food you love, a bowl of ice cream or a lush organic strawberry. All you have to do is imagine it and your tongue will already start to salivate whether or not you have any ice cream or strawberries around.

It's unlikely that the words "**predictive analytics**" spark the same level of salivation, but that's exactly what your brain is doing. Your brain is predicting and pre-sensing enjoyment based on a stored memory. It is acting on a model you've stored for a certain food. This model classifies data and our brain predicts our responses from the patterns within that data. From these patterns your brain starts to develop rules. The rules become the basis for your brain's automatic behaviours and responses.

Data

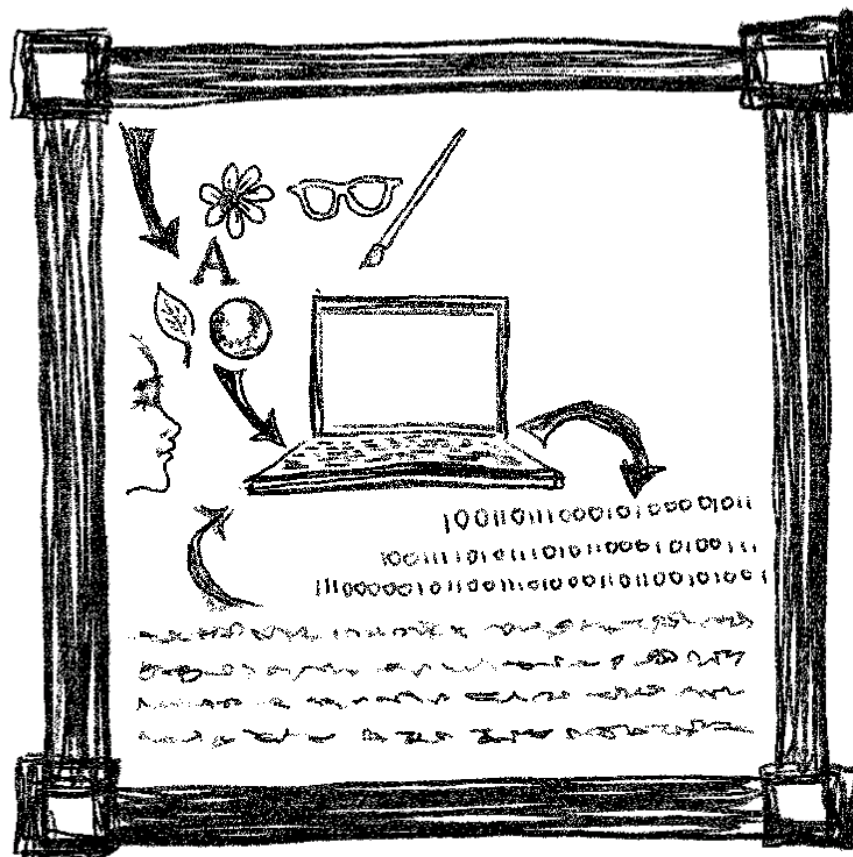
In the same way you have gone about your life building predictions about things you like, computers all over the world are making predictions about your behaviour based on your online activity. On average everyday, humans around the world:

- Use Google 4 million times.
- View 4.5 million videos on YouTube.
- Exchange 188 million emails.

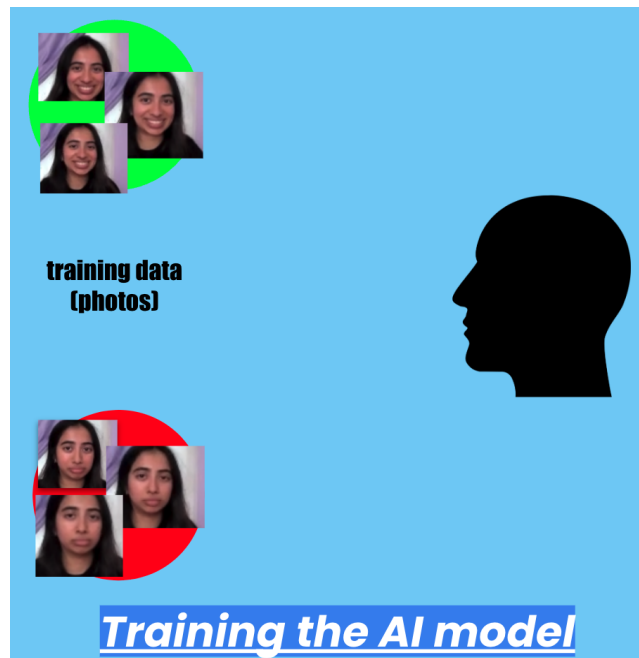
These exchanges produce a lot of data. Advances in chips and semi-conductors have greatly accelerated how easily and how fast computers can now work with large sets of data, which has been an important contributor to the development of machine learning. Computers sift through Big Data to make predictions about what you like or don't like, and often will respond to those predictions with suggestions for stuff you might like to buy, watch, or consume.

Let's look a little more closely at this process:

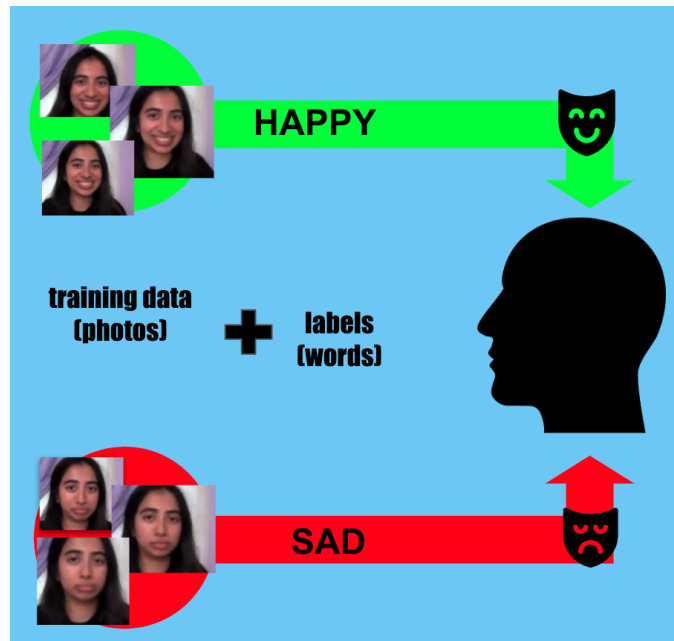
Training AI Algorithms



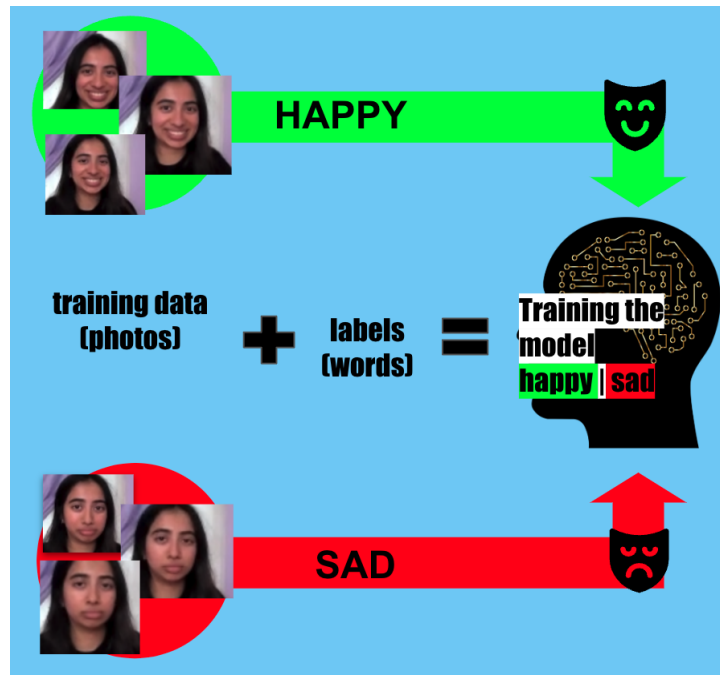
To better understand how AI can predict what you will or won't like, let's look at the process by which an AI might learn what kinds of things make you happy or sad.



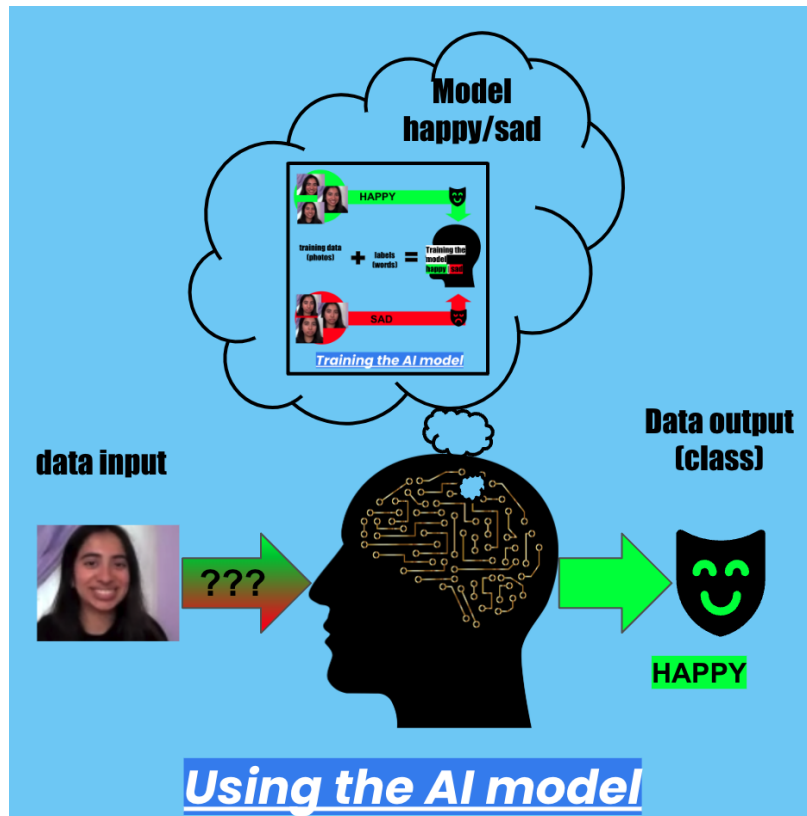
Before the AI can start predicting your responses, it needs to label the data that it is collecting. In this case, AI is being trained on screenshots of a face to distinguish happy faces from sad faces. It will need a fair number, let's say a minimum of 30, to see the common features of a "happy" face.



AI has no inherent understanding of happiness or sadness. It can only divide things into categories that have been labelled with variables, like the words “happy” and “sad.” AI doesn’t care what the labels are (because it is not a person with the ability to care), it just needs enough data that it can sort objects in the patterns of data sets that we give it. With enough examples of smiling faces that are labeled “happy,” and downcast faces labelled “sad,” the AI can create a working model to test how well it does with much larger sets of data. While it may be able to create a preliminary model with only 30 examples, to strengthen and refine the model it will benefit from much larger and more diverse examples.



With this model, the AI can then “output” its guess of what emotion the face is conveying and test whether or not its guess is correct. As AI takes in more data and tests its own predictions, it develops increasing accuracy because its choices are based on greater and greater numbers of real human examples. The algorithm that enables the AI to build its own set of rules is called a “neural network.”

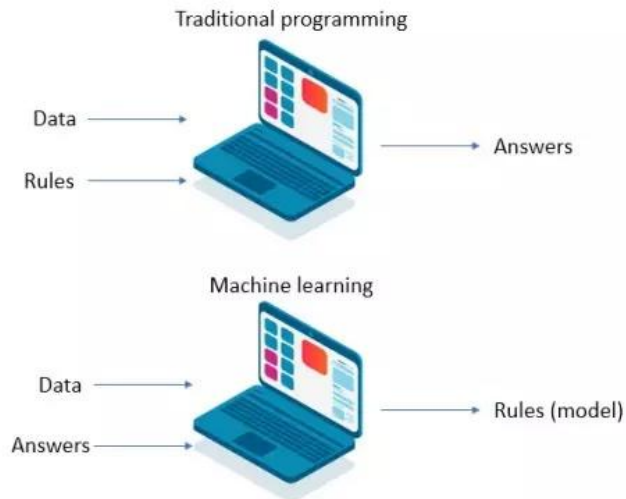


The important difference between machine learning algorithms, or neural networks, and the traditional algorithms used before advances in AI, is that in the past humans made the rules that the machines followed.

A traditional computer algorithm might have been a program that would provide pictures that a human had already labelled and then have the computer output answers based on this pre-labelled data.

In machine learning, we may label some large sets of pictures. But then we let the algorithm infer the similarities, and then test itself on whether it is correctly labelling new data.

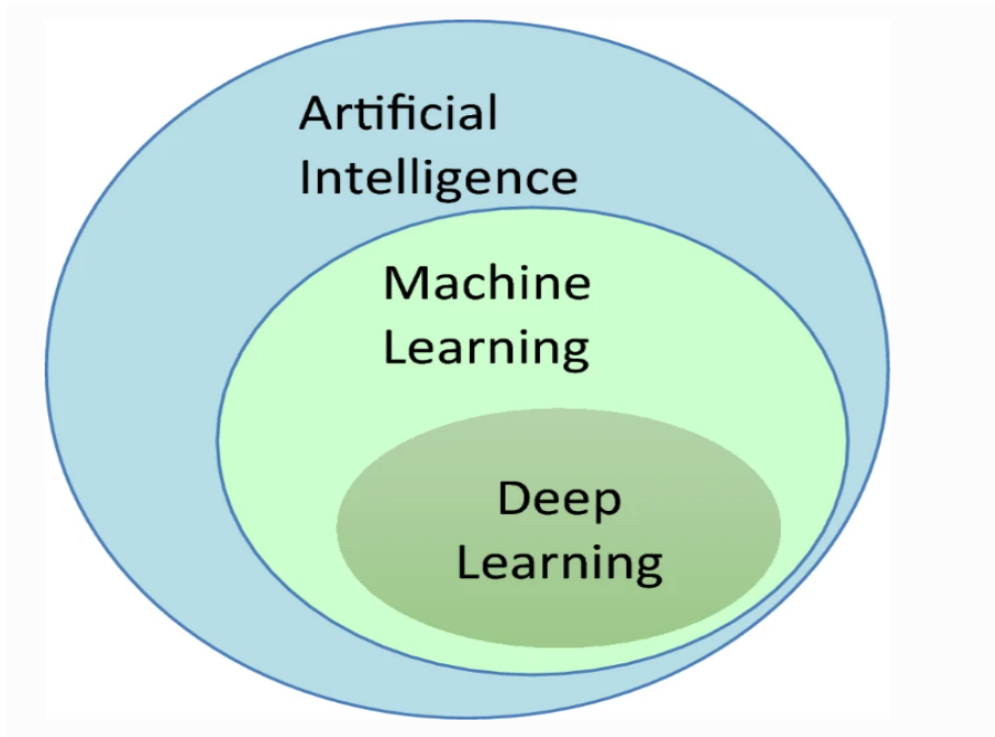
The more data the computer is trained on, and has learned from, the more the machine itself is making the rules from the model it has created.



This is important! It will always be worth keeping in mind that just because a computer algorithm can make a lot of rules, this doesn't mean that it should be making all the rules. It especially does not mean that all the rules are always right!

The kind of learning that AI does is still what we call "narrow" learning. Unlike humans, AI is not equipped to put its learning into context the way that humans do.

The more data driven AI algorithms are, the more they enter the state that is referred to as "deep learning". Here the layers of neural networks are many and complex. They are less transparent, and more difficult to explain. If an AI has a tragic meeting with a kangaroo, it is probably because there hasn't been enough human involvement in the training and shaping of the algorithm for all the different situations a human might encounter.



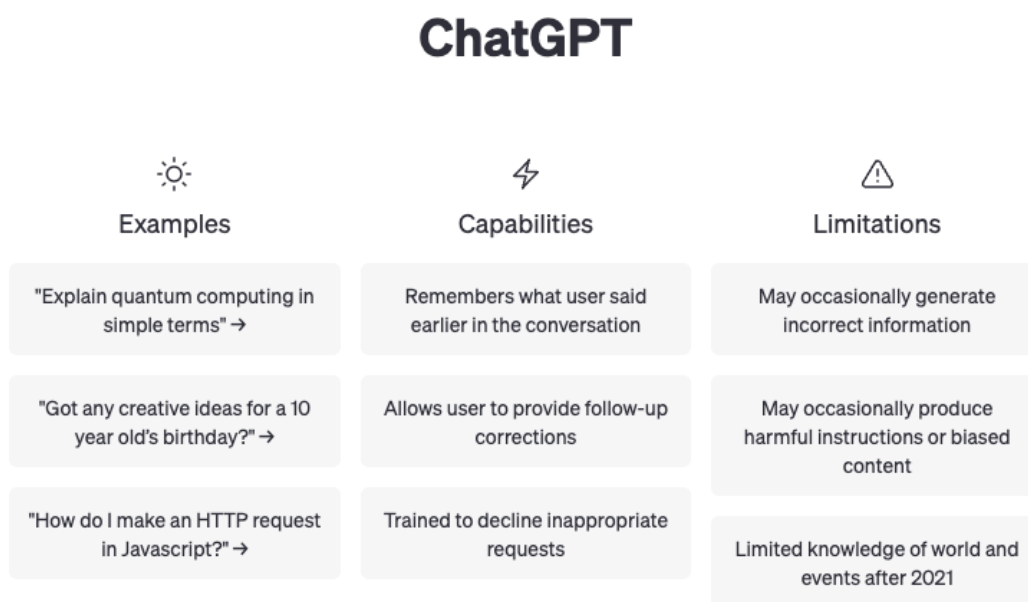
Let's go back to our "happy/sad," model. There are plenty of reasons why AI might still get this wrong, or make correlations that are irrelevant. Perhaps, for example, it was trained on thousands of faces, but those faces were only caucasian. Or perhaps its users were from a culture that was more likely to respond with unhappiness to something that produced happiness in a dominant culture.

Humans can and should continue to influence the rules that machine learning operates from. The trick to optimal use of machine learning is knowing when we want machines to make the rules, and when we don't. We might, for example, benefit from hybrid systems that use machine learning for low risk automated learning, but bring in humans to ensure that data is both diverse and well balanced for important decisions where the risks are higher.

One way to clarify decisions like these is to use "model cards." These are documents that AI creators can use to bring transparency to how an AI model is intended to be

used, its limitations and warnings, what data it was trained on and what labels and features were used to train the data.

Here, for example, is a model card created for Chat GPT, the impressively proficient chatbot. This model card was introduced a few months after the first wave of users realized that ChatGPT wasn't always as proficient as it first seemed:



Invisible stumbling blocks

Model cards are one of the ways that an AI creator can help a user scope whether this AI is right for the problem the user wants to solve, or the artifact the user wants to create.

But not every AI-driven system a citizen living in the digital age will encounter comes with a model card. One of the most pervasive challenges of living in the AI era is to be aware of the ways that AI functions that are invisible or obscure to you. When a

streaming platform makes recommendations to you, or when ads for something you recently mentioned in an email or on social media magically appear, you know that AI has been tracking and predicting your behaviour.

Every year AI gets better at predicting what humans will do or say. If you've tried out a conversational AI tool like ChatGPT, you've seen how effectively AI can cut and paste the sentences it believes would make good answers to your questions.

You may already have seen firsthand how ChatGPT can fail. It may provide an answer that sounds very credible, but is fully or partially false. Often the biggest errors are not created by the AI, but by the humans using the AI. The biggest mistake we can make is in assuming that just because AI is doing a great job of mimicking humans that it is actually human.

One of our most universal human tendencies is anthropomorphism, the tendency to project our human intelligence onto other species and even objects. We often believe our pets are smarter than they actually are, so it is easy for us to assume that the AI in our lives have human, or even magical powers.

But in all likelihood there will always be things that humans do better than AI, especially when it comes to understanding context and what is called "broad" learning. We are still better at being able to make the kind of visual and spatial connections that help us judge how close or far a Kangaroo is to our car, even if we're not Australian!

As chatbots become more common in the world, it will become more and more important to know when you are interacting with one.

We don't know whether AI will replace humans. It doesn't seem likely, no matter how well a machine can be programmed to mimic humans. It is certainly safe to say that AI is challenging humans to become smarter and even more skilled by helping humans get better at the repetitive learning tasks that we are not as well suited for.

But this also puts some pressure on us to demonstrate our superior ability to do the kind of work that AI can't do, such as think critically about situations, make decisions about what we value, understand why we value those things, and care about the welfare of others.

What machine learning teaches us about our own learning

One of the greatest opportunities presented by AI is that it challenges us to better understand the important differences between how machines learn and how humans learn. As we become better at knowing and judging these differences we will become better at making decisions about how to use AI ethically and effectively.

Learning about AI can teach humans several things about their own intelligence:

1. **Limitations:** One of the most significant lessons we can learn from AI is that there are inherent limitations to our own intelligence. AI models can perform complex computations at lightning speeds and can remember vast amounts of data with ease, which can highlight the limits of our own cognitive abilities. This can help us identify where we need to improve and where we need to rely on technology to help us.
2. **Problem-solving:** AI algorithms are designed to solve specific problems, and studying how they do so can teach us new ways of problem-solving. By understanding how AI approaches problems, we can apply similar techniques to solve our own challenges, whether in business, science, or everyday life.
3. **Bias:** AI algorithms are only as unbiased as the data that is fed into them. If there is bias in the data, it will be reflected in the results. By studying AI, we can learn about the risks of bias and how to mitigate it in our own decision-making.
4. **Creativity:** Although AI algorithms are becoming increasingly sophisticated, they still lack the creativity and imagination that humans possess. This can serve as a reminder of the unique qualities that make us human, such as our creativity, intuition, and emotional intelligence.

Overall, studying AI can help us understand both the strengths and limitations of our own intelligence, and can inspire us to find new and innovative ways to solve problems and create value.

E is for excitement and the end of this primer (but only the beginning of your journey)

While the progress of AI seems to be accelerating, the road to using it to its full potential will likely be filled with debates over its uses and dangers and a growing awareness of the competencies we will need to master to understand it better. We will need to know more than simply how to interact with AI, and need to develop basic proficiency in emerging literacies like algorithm and data literacy, modelling and machine learning.

Excitement about AI may often transform into anxiety, even disillusionment and despair. Some of the evolution of AI will be in our control, some of it won't. For this reason it's important to ground our excitement in the joy of learning.

In the coming years there will be so much to learn about AI, what it can do and can't do. There will also be so much to learn about the human species, what it can't do, but also things we can now do that we're once beyond our ability to imagine.

As we look towards the future, to maintain our excitement it might be useful to look to the past and think of the seven generations principle employed by the Haudasanee (Iroquois) nation to sustain one of the oldest democracies in the world.

If we design and use AI with an intention that respects all that we owe to the actions, talents and skills of the generations that have come before us, and all that we aspire for the generations who will come after us, we can at least know that we have done what we can to ensure that AI will become and remain a vital part of a thriving planet in which all species have value.

Resources

[The Algorithm and Data Literacy Project](#)

["AI and Creative Learning: Concerns, Opportunities, and Choices"](#) - Mitch Resnick

[Machine Learning in a Day](#) - Apps for Good

[The Actua AI Handbook](#)

[Dancing With AI](#)

[A quick reference guide to artificial intelligence for teachers](#)