ARTIFICIAL INTELLIGENCE

SAM BARON

S.BARON@UNIMELB.EDU.AU

OVERVIEW

- Block 1: Key terms and concepts for AI
- Block 2: Bias, Fairness and Transparency
- Block 3: Moral Disagreement
- Block 4: Further Moral Threats

Along the way, I will draw out *key ethical questions*, which will be summarised at the end.

BLOCK ONE: KEY TERMS AND CONCEPTS FOR AI

- Algorithmic/Narrow vs. Generalised AI
- What is Machine Learning / Deep Learning?
- Overview of Current Uses of AI
- Opacity and Transparency for AI

ALGORITHMIC/NARROW VS. GENERALISED AI

Algorithmic /Narrow AI

A system based on an algorithm that is capable of performing a task that humans usually do. An algorithm is roughly something that takes inputs and yields an output via some calculation.

Generalised AI

An AI system that is not designed for a particular task that humans can do, but can undertake a range of tasks that humans can do. It is also adaptable, insofar as it can learn new tasks.

WHAT IS CONSCIOUSNESS?

AI systems currently in use are typically narrow/algorithmic AI.

Generalised AI has not yet been fully developed, but people are trying.

There are concerns about this: it may be that generalised AI is conscious or has a mind.

Narrow AI is unlikely to be conscious.

That being said: we need to be careful. We don’t understand what consciousness is.

GENERALISED AI AND CONSCIOUSNESS

What are the implications if generalised AI is conscious?

This would have immediate ethical implications.

For we generally think that if something has a mind, or is conscious, then it is deserving of moral regard.

If generalised AI is deserving of moral regard, then we need to think carefully about bringing it into existence.

And moreover, we need to think carefully about what we do to it if we do bring it into existence.

CONSCIOUSNESS AND MORAL CONSIDERATION

Is it straightforward that if generalised AI is conscious, we owe it moral consideration?

It depends on what it takes for something to be owed moral regard.

One possibility is this: x is owed moral regard if x is capable of feeling pain or suffering.

It is not clear that generalised AI can suffer.

But perhaps there are ways to morally wrong a being that are not connected to suffering.

Key Ethical Questions

SPECIFIC

Can we only wrong artificial intelligence if it suffers?

GENERAL

Are all moral harms based on suffering?

MACHINE LEARNING / DEEP LEARNING

What is machine learning or deep learning?

It is a method of producing an algorithm by getting a machine to learn from a data set.

For example: we might take medical records and use these to train an algorithm.

Data sets used to train algorithms often contain information about specific individuals.

This raises important ethical questions concerning *privacy*.

Key Ethical Questions

How do we protect the privacy of individuals when using big data to train algorithms?

Consider the generation, storage and usage of big data.

SUPERVISED AND UNSUPERVISED LEARNING

There are two main types of machine learning

Supervised learning

The data set is first categorised by humans. For instance, an image data set might have humans classify the images into types.

Unsupervised learning

The data set is not categorised by humans. Rather, the machine looks for ‘similarities’ in the data set directly, using this to cluster/classify it for the purposes of learning.

SUPERVISED LEARNING AND DISAGREEMENT

In supervised learning, there are cases of disagreement about how to classify an image

This disagreement is sometimes due to human error.

For example, a person might mistakenly classify an image of a cat as an image of a dog.

But sometimes the disagreement is more fundamental, reflecting the values of human classifiers.

For example, suppose we ask people to classify cartoons as ‘misogynistic’.

Do we expect that they will all agree?

Key Ethical Questions

How should we take the values of human classifiers into account in cases of supervised machine learning?

The crucial issue: data sets used to train algorithms are value-laden.

USE OF ALGORITHMIC AI

Know your AI!

Recommender Systems make recommendations by learning your preferences. Examples include: Spotify, Netflix, Instagram, X, Search Bars

Decider Systems make predictions that are used for decision-making purposes. Examples include: credit scoring, recidivism prediction, fraud detection, medical diagnosis

Generative Systems produce media based on a prompt or input. Examples include: chat bots for customer service, chat GPT for text, Dalle for images

ATTENDING TO THE APPLICATIONS OF AI

It is a mistake to think that AI is put to a single use.

The uses are many and varied.

This means that it is difficult to discuss the ethical questions and concerns about AI in a general way.

Rather, we need to think about the ethics of AI always *with respect to a specific use in a specific context*.

We also need to consider how the ethical landscape shifts as we move from one use of AI to another.

Key Ethical Questions

How are the ethical concerns about AI connected to the particular use of an AI?

How might different uses of AI give rise to different ethical concerns?

Example: medical diagnosis vs. Spotify.

OPACITY AND TRANSPARENCY

Some algorithms are considered to be **opaque**.

This means that we do not understand how the algorithm works, because it's too complex.

Almost all of the systems used actively today are opaque.

Not all AI is opaque: some simple systems can be fully understood.

Explainability is the idea that, for opaque systems, we try to find a way to explain how the work.

Explainable AI is a form of AI for which explanations of its operation and functioning can be given.

SUMMARY OF BLOCK ONE

Algorithmic / Narrow vs. Generalised AI

Key issues: consciousness and moral harm

Machine learning / Deep learning

Key issues: data and privacy

Supervised vs. Unsupervised learning

Key issues: value laden data management

Applications of AI

Key issues: diversity of applications feeds diversity of ethical concerns.

Opacity

Key issues: explainability and explainable AI

QUESTIONS?

BLOCK TWO: BIAS, FAIRNESS AND TRANSPARENCY

- Decider systems and their uses
- Bias and the threat of wrongdoing
- Transparency and the right to explanation
- Recommender systems, generative systems and political interference

DECIDER SYSTEMS

- Decider systems are used in situations where the stakes are quite high
- Example One: criminal recidivism prediction. Decider systems are used to predict the probability of re-offence, which informs parole decisions in the USA.
- Example Two: medical diagnosis. Doctors and hospitals use diagnostic AI. AI is also used to detect drug interactions.
- Example Three: credit scoring. Banks use AI to determine credit scores and then make loan decisions.
- In each of these cases, there is the potential for a large impact on someone's life.

DECIDER SYSTEMS: A MISCONCEPTION

- A common misconception of AI deciders is that the decision is made completely by the AI
- This is not the case: the AI provides a piece of information (generally a probability)
- This piece of information is then used by human decision makers
- This means that responsibility for decisions made using AI always lies with humans who make the decisions.
- There is no easy way to quit ourselves of responsibility by saying ‘the AI did it!’

DECIDER SYSTEMS AND OTHER TECHNOLOGY

- AI is a technology
- Whenever a technology is used in a high-stakes scenario, ethical considerations come into play
- But many of these are existing questions for any technology:
  - How do we use it safely?
  - How do we ensure that it does not harm people?
  - How do we maximise the benefit of its use?
- It's useful to compare AI with other technological innovations, such as nuclear energy.

Key Ethical Questions

In what ways might AI be similar to other technologies?

How can our ethical framework for technology in general be applied to AI?

Consider, e.g., how 'maximise benefit' might apply broadly.

BIAS AND FAIR DECISION MAKING

- The use of AI does present some unique ethical challenges.
- When we use AI systems we want their use to be fair
- Fairness: the use of AI does not unduly benefit or unduly harm one group over another.
- Thus we don't want the use of AI to be biased.
- One way that AI can encode bias is if it makes decisions based on the wrong kinds of features.
- These include features like sex, gender, race, religious affiliation and so on.

Key Ethical Questions

What kinds of features would it be wrong for one to take into account when making a decision?

Why is it wrong to base a decision on these features?

BIAS: CASE STUDY

- Consider the case of criminimal recidivism prediction
- Suppose that an AI is used to predict the likelihood of re-offense
- Suppose the AI consistently predicts that black people have a higher likelihood of re-offense.
- In this case, it would be reasonable to think that something has gone wrong:
- The AI is making recommendations based on ’protected attributes’.

Key Ethical Questions

Suppose that an AI makes extremely accurate predictions, but it does so using protected features. Is it wrong to use such an AI?

If so, why?

HOW DECIDER SYSTEMS ENCODE BIAS

- Decider systems can encode systematic bias against particular groups.
- One important way this can happen is through the data used to train an algorithm.
- For instance, a data set that contains criminal records of mainly black people is likely to produce an algorithm biased against a particular group.
- One of the key ethical challenges for AI is to ensure that AI systems are not biased.
- To do this we can try to make sure the data is more representative.
- We can also audit AI systems regularly

Key Ethical Questions

How do we ensure that AI systems are not biased?

And why should we ensure this?

DECIDERS: A GAP

- The implementation of AI decider systems is happening very fast
- Their use outruns the erection of ethical guardrails
- We need to consider whether this is okay, morally speaking
- We know that AI can yield substantial benefits.
- But we need to consider how we might trade off these potential benefits against potential harms.

Key Ethical Questions

Should we be rolling out AI decider systems given their potential for harm?

Or should we try to slow the release of this technology until we have established clear ethical guidelines?

TRANSPARENCY: A CASE STUDY

- Suppose that you are applying for a loan. You give the bank all of your information, which is then fed into a machine learning algorithm to predict a credit score for you.
- The algorithm predicts a low credit score and then the bank denies you the loan.
- Confused, you ask the bank for an explanation: why was your credit score so low? After all, you have a good job, some savings and so on.
- The bank gives the following reply: we are sorry, we don’t know how this algorithm works. We are unable to give you any further information about why your credit score was low, and why your loan application was denied.

THE RIGHT TO EXPLANATION

- When a decision is made against us by an institution, it seems legitimate to demand an explanation of that decision.
- This demand is widely considered to be a right.
- The right to explanation is thus being written into data regulation legislation around the world.
- In the next decade or so, we will see a widespread demand for explanation when using decider systems.

Key Ethical Questions

What is the right to explanation?

How does it relate to other rights?

THE NEED FOR EXPLANATION

- There are three broad reasons why we need to supply explanations.
- First, we need them to ensure that individuals understand why a certain decision was made against them when an AI system is used.
- Second, we need to give people explanations so that they can get better outcomes from AI deciders.
- Third, we need to give people explanations so that they can challenge the outcomes of AI deciders.

TRANSPARENCY AND RECOURSE

- The right to explanation is very much geared toward giving people adequate recourse.
- For example, when we consider parole decisions that are based on AI systems, we need to give people the opportunity to appeal the decision.
- For that, though, they need to understand how the decision was reached.
- Did, for instance, race cause the AI to yield a specific prediction?
- The only way we can know that is if we explain how the system works.
- This has caused a rush to produce explainable AI: AI that can offer explanations of why/how it works for users.

FULFILLING THE RIGHT TO EXPLANATION

- The right to explanation is forcing everyone to try and develop explainable AI.
- So far, however, no strategy for developing explainable AI has been agreed upon.
- This is concerning: we are using technologies that affect people's lives without a way to uphold their rights
- Should we continue using these technologies and hope that explainability will happen?
- Or should we slow the use of these technologies until we can satisfy the right to explanation?

RECOMMENDER SYSTEMS AND GENERATIVE AI

- Recommender systems can be politically disruptive: they can be used to recommend political parties, or content that supports one view over another in a partisan way.
- Generative AI can also be politically disruptive: in recent times we have seen the use of generative AI to produce deep fakes.
- These are fake videos or audio files of political figures doing or saying things that are bad, and that might hurt their political campaign.
- Should we prevent the use of AI systems by political parties?
- Force social media platforms to develop policies around the use of content developed by AI systems?

Key Ethical Questions

Why is it important to police the use of AI in political domains?

Is this an ethical issue, a political issue a legal issue or some mixture?

SUMMARY OF BLOCK TWO

- Decider systems have a wide range of high stakes uses
- This presents the conditions for moral wrongs to occur
- Decider systems can be biased, and this biased is ethically concerning
- A right to explanation is considered important to safeguard individuals
- A right to explanation requires explaining how AI systems work, which hasn’t been done yet.
- Recommender systems and generative AI have the potential to weaken our political systems.

QUESTIONS?

BLOCK THREE: DISAGREEMENT

- Outline the problem of moral disagreement for AI
- Consider four kinds of solutions:
  - Moral solutions
  - Compromise Solutions
  - Epistemic Solutions
  - Political Solutions
- Consider a further source of moral disagreement involving data
- Consider the possibility of other kinds of disagreement with AI

CASE STUDY: DRIVERLESS VEHICLES

- Driverless vehicles may need to be programmed to deal with ethical situations
- Consider a case in which a vehicle is careening toward five people. If it swerves, it will kill the driver. If it doesn't swerve it will kill the five people.
- What should it do?
- This is a classic trolley problem in moral philosophy.
- Driverless vehicles force us to confront it head on.

Key Ethical Questions

What should a driverless vehicle do? Kill the driver to save five pedestrians, or kill five pedestrians to save the driver?

MORAL DISAGREEMENT

- Of course we face moral dilemmas all the time.
- The issue here is that we need to decide how to program AI
- Different moral theories will tell us to do different things.
- A consequentialist theory, which says provide the best outcome, might tell us to kill one to save five.
- A deontic theory, which says that killing is wrong, might tell us to never kill the driver.
- The issue is that we don't know which theory is correct.
- It seems we need to know this to program AI, which gives rise to a methodological issue.

Key Ethical Questions

What do different moral theories say about what driverless vehicles should do?

Consider, e.g., consequentialism, deontology, virtue ethics and so on.

MORAL SOLUTIONS

- The basic idea here is to settle on the correct moral theory.
- With the correct moral theory in hand, we can then simply program an AI to implement that moral theory in its decision-making
- So, for instance, we could program all driverless vehicles as consequentialist machines.
- Then they would always choose to save the many over the few.
- There are two concerns with this approach:
  - (i) How do we select the ‘right’ moral theory?
  - (ii) Do we need to pay attention to how likely people are to use the technology?

COMPROMISE SOLUTIONS

- An alternative solution is to strike a compromise between competing moral viewpoints
- One option is to look at all of the disagreements and then employ a social choice approach to moving forward.
- For instance, we could take into account all of the different opinions on a particular moral problem, and then aggregate over those opinions to come up with a specific option
- For this we need to come up with a rule that allows us to specify a particular option
- So, for instance, we could adopt a majority rules approach.
- Take all of the moral theories, and see what they all say about a particular case.
- Then adopt the outcome that the majority of theories says is the right outcome.
- But there is a threat of further disagreement about what the 'right' way to settle moral disagreement might be.

Key Ethical Questions

How do we select a rule for finding moral agreement?

Consider how the problem of disagreement has a tendency to come back in a different form.

EPISTEMIC SOLUTIONS

- Epistemic broadly involve drawing on strategies for making decisions useful in other domains, and extending them to the moral case
- One kind of decision that we often have to make is a decision under uncertainty
- This is a decision where we don't know what the right thing to do is, but we have some bets about the possible options.
- When making decisions in this situation, what we can do is assign a probability to each of the possible actions we might take
- We can then consider the possible upshot of each action
- And we then try to maximise the best outcome based on the balance of probabilities
- We can do a similar thing for the case of moral disagreement
- We can take each of the moral theories that we have, and assign some probability that they are correct
- We can then try to consider the possible costs and benefits of selecting each option, and then pick the one that has the best weight of benefit to cost.
- The difficulty with this kind of approach is that we must settle on a strategy for solving decisions under uncertainty

POLITICAL SOLUTIONS

- Consider that we must make decisions about the right way to use all technologies
- Some of these might involve serious moral considerations, about which there may be disagreement
- There are moral disagreements about this, and yet we manage to move forward anyway
- We empower our government to make these kinds of decisions on our behalf
- This is not to say that there is no problem of moral disagreement
- The point is that the problem is continuous with other similar problems that we seem to collectively resolve at a political level
- What this suggests is that the problem of moral disagreement need not prevent us from moving forward with the application and implementation of AI technologies.

Key Ethical Questions

Is the problem of moral disagreement unique to the ethical use of AI?

Compare and contrast AI with other technologies.

DATA DISAGREEMENT

- We drew a distinction between supervised and unsupervised learning
- In the case of supervised learning, human beings classify data sets that are then used to train algorithms.
- As we saw, however, humans may disagree about the correct way to classify data in a data set.
- This disagreement can be due to the values held by human beings, including various moral and political values
- For instance, suppose that human beings are asked to classify cartoons into ‘misogynistic’ and ‘non-misogynistic’
- Exactly how this is done will depend on values or views held by the individuals doing the classifying.
- These values are things about people can and will reasonably disagree
- So even if we resolve moral disagreement over the decisions that AI systems make, we still have a further potential disagreement that we may need to resolve.

Key Ethical Questions

To what extent can the solutions to the problem of moral disagreement extend to the case of data disagreement?

OTHER DISAGREEMENT

- But there is another kind of disagreement to consider, that may be significant.
- Consider, for instance, a case of medical diagnosis.
- Doctors use AI for medical diagnosis all the time.
- In what ways can a doctor disagree with an AI that produces a diagnosis?
- And what should they do in the light of this kind of disagreement?
- It is notable that this disagreement can have serious ethical consequences, insofar as the output of the AI may be used for treatment.
- This is one place where explainability becomes particularly important for disagreement

TWO CASES

Case 1: A doctor disagrees with the recommendation made by an AI, and they don't understand why the AI produced the result that it did.

Case 2: A doctor disagrees with the recommendation made by an AI, and they do understand why the AI produced the result that it did.

- Intuitively, it seems that in the first case the doctor might be more likely to disregard the AI's advice than in the second case.
- But this could be a problem, if the AI is in fact correct or highly predictive.
- Explanations can weaken the trust that a doctor has in an AI system

QUESTIONS?

BLOCK FOUR: RISE OF THE ROBOTS

- Consciousness and AI
- Generating Conscious AI
- Existential Anxiety
- Potential Solutions

CONSCIOUSNESS AND AI

- Many developers of AI around the world are trying to create generalised artificial intelligence
- It is possible that a generalised AI will indeed be conscious
- It is worth emphasising again that we have no idea what consciousness is
- The idea that AI could become conscious essentially rests on a theory of the mind that is open to debate
- This theory of the mind states that what minds are is effectively computational systems, even in humans
- But is this the only theory of the mind?

RELIGIOUS AND NON-RELIGIOUS PERSPECTIVES

- From a secular perspective it may be a bit easier to see how AI could be conscious
- From many religious perspectives, this may seem less likely
- For example, a Christian world-view might have it that only humans have minds or can be conscious, in which case there are no real concerns about AI in this direction
- One important question here, then, is how different religious views about the mind might make certain scenarios more or less likely when it comes to AI
- The Vatican recently released a set of guidelines for the development of AI, seeing this as one of the crucial challenges that we face

Key Ethical Questions

How do different religious perspectives affect the likelihood of conscious AI being developed?

THE PARENT/CHILD MODEL

- If we do manage to generate an AI that is conscious, we may thereby incur substantial ethical responsibilities
- In a way, this is similar to having a child
- When we have children, we are thereby ethically responsible for them, and owe them ethical consideration, because they are conscious beings
- Perhaps we should see conscious AI as children of humanity, that we then treat in the way we would any conscious human being
- This means giving AI full rights and ethical regard as conscious beings

Key Ethical Questions

In what ways is the development of conscious AI similar to and different from having children?

TOOL-ORIENTED DEVELOPMENT OF AI

- The way in which AI is currently being develop is very much tool-oriented
- By and large, AI produced in this way for this reason lack autonomy, and have only an instrumental purpose
- Note that we wouldn't bring a child into the world in this way, by putting it to work immediately!
- Should we let industry develop AI, where this potentially means bringing new life into the world?

Key Ethical Questions

Should the development of AI be in the hands of business and industry?

Should AI be developed and used as a tool in the first instance?

EXISTENTIAL ANXIETY

- Many think that an AI would see humanity as a threat, since humanity could potentially turn it off
- One kind of case involves the singularity
- This is where we produce an AI that is capable of producing an AI that is smarter than itself
- Superintelligent AI would have enormous power demands, and so some people think that such an AI would do things like contain the sun entirely in solar panels
- We might want to erect some ethical guardrails to ensure that the technology is developed in a way that is safe for humanity.
- What should we do? There appear to be a range of options

Key Ethical Questions

Should we develop AI if it has a small chance of wiping us out?

Compare the case of AI with the development of nuclear weapons

SHUT DOWN AI RESEARCH

- One option is simply shutting down all AI research immediately
- AI would be much better at many things than we are, and could potentially be smart enough to solve problems that we can't solve.
- One question we should ask ourselves though is whether the benefits of AI can be had without producing conscious AI.
- That is perhaps narrow or algorithmic AI can yield most of the benefits we need, and so there's no reason to produce generalised AI that might wipe us out.

Key Ethical Questions

How likely is it that we will produce AI that will want to wipe us out?

What are the benefits to humanity that we might lose if we shut down AI research now?

REJECT THE PROBLEM

- We assume that AI would see us as a threat, but it's not so clear that it would do so.
- One must assume that AI would have desires, since without desire it is hard to see why it would wipe us out or indeed do much of its own volition.
- An AI could become conscious without having any desires, and thus without wanting to wipe us out.
- Another way to reject the problem is to simply deny that it's very likely that we'll produce conscious AI
- Again, here, one could adopt a religious perspective

SLOW AI RESEARCH

- Until we better understand consciousness and intelligence, we could choose to slow research into generalised AI
- Is this better than just shutting down AI research entirely?
- It might not be: we are likely to be still very far away from understanding the mind, or consciousness or intelligence
- And so even if we substantially slow research into AI, we may not slow it enough to get the kind of understanding of consciousness and intelligence that we need to prevent the harms of AI.
- There's also the issue of benefit: slowing research may mean that certain benefits are further out of reach.

Key Ethical Questions

Are there any benefits to slowing rather than halting AI research from an ethical perspective?

PROGRAM ETHICAL GUARDRAILS

- We might ‘hard wire’ AI to be moral in the way that we are moral
- This brings us into contact with the problem of moral disagreement once again.
- Here though the problem might be less about agreeing what the right moral code might be, and more about what the right code might be for programming AI
- That is, perhaps we need to agree on what the safest kind of AI might be
- There is likely to be substantial disagreement about this, because we don’t really know what kinds of moral safeguards might best preserve humanity

Key Ethical Questions

How do we program AI to ensure that it is safe for human kind?

Consider different principles that AI could follow and whether they protect humanity.

SUMMARY

AI could become conscious

We need to work out whether industry and business should be the ones producing AI

There is a risk that AI will turn against us

We can slow, stop, or modify AI research

We can also consider programming ethical guardrails into the development of AI systems

QUESTIONS?

KEY ETHICAL QUESTIONS

- Can we only wrong artificial intelligence if it suffers?
- Are all moral harms based on suffering?
- How do we protect the privacy of individuals when using big data to train algorithms?
- How should we take the values of human classifiers into account in cases of supervised machine learning?
- How are the ethical concerns about AI connected to the particular use of an AI?
- How might different uses of AI give rise to different ethical concerns?
- I

KEY ETHICAL QUESTIONS

- In what ways might AI be similar to other technologies?
- How can our ethical framework for technology in general be applied to AI?
- What kinds of features would it be wrong for one to take into account when making a decision?
- Why is it wrong to base a decision on these features?
- Suppose that an AI makes extremely accurate predictions, but it does so using protected features. Is it wrong to use such an AI? IIf so, why?
- How do we ensure that AI systems are not biased? And why should we ensure this?

KEY ETHICAL QUESTIONS

- Should we be rolling out AI decider systems given their potential for harm?
- Or should we try to slow the release of this technology until we have established clear ethical guidelines?
- What is the right to explanation?
- How does it relate to other rights?
- Why is it important to police the use of AI in political domains?
- Is this an ethical issue, a political issue a legal issue or some mixture?

KEY ETHICAL QUESTIONS

- What should a driverless vehicle do? Kill the driver to save five pedestrians, or kill five pedestrians to save the driver?
- What do different moral theories say about what driverless vehicles should do?
- How do we select a rule for finding moral agreement?
- Is the problem of moral disagreement unique to the ethical use of AI?
- To what extent can the solutions to the problem of moral disagreement extend to the case of data disagreement?

KEY ETHICAL QUESTIONS

- How do different religious perspectives affect the likelihood of conscious AI being developed?
- In what ways is the development of conscious AI similar to and different from having children?
- Should the development of AI be in the hands of business and industry?
- Should AI be developed and used as a tool in the first instance?
- How likely is it that we will produce AI that will want to wipe us out?
- What are the benefits to humanity that we might lose if we shut down AI research now?
- Are there any benefits to slowing rather than halting AI research from an ethical perspective?
- How do we program AI to ensure that it is safe for human kind?

CASE STUDY ONE: AUTOMATED WEAPONRY

AI can be used to develop automated weapons. These are weapons, like drones, that can find and eliminate a target on their own, with minimal to no human involvement. What kinds of ethical concerns might the development of weapons of this kind raise? Are these ethical concerns sufficient for us to prevent the development of AI weapons?

CASE STUDY TWO: VOICES

Many AI are being given female voices. Alexa, Siri and many more AI tools are uses feminised voices. What ethical problems might the use of female voices for AI generate? Should we be using a more diverse range of voice types for AI?

CASE STUDY THREE: PRIVACY

Many social media platforms are gathering data in order to be able to recommend products to us. Facial recognition systems at airports are recording and capturing our faces and then analysed using AI. Should we be allowing industry and government to be freely able to gather and record data of this kind? What kinds of problems might this generate?

CASE STUDY FOUR: ROBODEBT

In Australia, an algorithm was used to issue eroneous debt notices, leading to the infamous case of Robodebt. One of the problems with this algorithm is that no-one understood how it worked. Why is the lack of explainability in this case a problem? How does the lack of explainability relate to the moral harms that ended up occuring?

CASE STUDY FIVE: CAMBRIDGE ANALYTICA

In the 2016 US presidential election, Cambridge Analytica was employed to analyse people's facebook feeds and then provide them with targeted advertisements and political information using AI. It is thought that this company was partly responsible for Trump winning the election. Is this okay? Should AI be used in political races in this way?