In this week, we will look at the FATE (Fairness, Accountability, Transparency, and Ethics) issues in AI and responsible data science.

Workshop Case Study

UBER EATS: delivering food while it's hot (extracted from [17 Data Science Applications & Examples](https://builtin.com/data-science/data-science-applications-examples) (https://builtin.com/data-science/data-science-applications-examples))

The data scientists at Uber Eats, Uber's food-delivery app, have a fairly simple goal: getting hot food delivered quickly. Making that happen across the country, though, takes machine learning, advanced statistical modelling and staff meteorologists. In order to optimise the full delivery process, the team has to predict how every possible variable — from storms to holiday rushes — will impact traffic and cooking time.

Discussion:

- What is the problem/challenge (problem definition/research question)?
- Why is it important (significance)?
- What is the AI/DS task? (see week 1 for the range of AI/DS tasks)
- What types of data do we need for optimising the delivery routes?
- Who are the users involved in this case study?
- Would there be potential ethical issues/concerns?

Amazon's Accent Recognition Technology

At the beginning of October 2018, Amazon was quietly issued a patent that would allow its virtual assistant Alexa to decipher a user's physical characteristics and emotional state based on their voice. Characteristics, or "voice features," like language accent, ethnic origin, emotion, gender, age, and background noise would be immediately extracted and tagged to the user's data file to help deliver more targeted advertising.

The algorithm would also consider a customer's physical location — based on their IP address, primary shipping address, and browser settings — to help determine their accent.

Discussion:

- Are there potential ethical concerns in relation to data privacy?
- Are there any concerns with regards to the intended use of this data and analysis?
- What rights do we have as consumers?

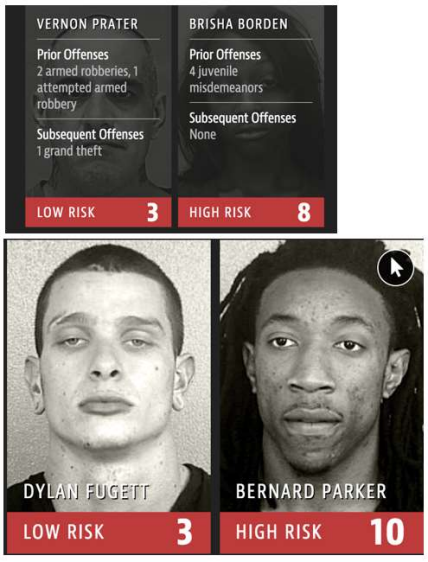
- Will AI models be limited if users don't want to provide data?
- Are there any other potential issues/concerns regarding FATE (Fairness, Accountability, Transparency, and Ethics)?

Pre-class reading:

Artificial intelligence (AI) is changing societies and economies around the world. However, despite all the benefits that AI bring us, the transformative power of AI, however, also comes with challenges, ranging from issues of privacy, bias, transparency, trust and security, to concerns about displacing jobs and exacerbating inequalities, etc. The more we study and develop artificial intelligence, the clearer it becomes that this massively powerful tool comes with a great deal of responsibility.

Fairness

Like humans, technologies can make mistakes, and often, they display unfair bias against people of color, gender, race, etc.. Take another example, Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), is a risk assessment algorithm powered by AI. COMPAS has been used in the United States to forecast which criminals are most likely to re-offend. However, it has been found to indirectly contain a strong racial bias, meaning black defendants are nearly twice as likely to be misclassified as higher risk than their white counterparts. This is not only unethical, but also unacceptable, for people to be disadvantaged in the application of these systems on a mass scale.



Source from [https://medium.com/thoughts-and-reflections/racial-bias-and-gender-bias-examples-in-ai-systems-7211e4c166a1 (https://medium.com/thoughts-and-reflections/racial-bias-and-gender-bias-examples-in-ai-systems-7211e4c166a1)](https://medium.com/thoughts-and-reflections/racial-bias-and-gender-bias-examples-in-ai-systems-7211e4c166a1)

Transparency and Explainability

Another potential problem that comes with AI is the lack of transparency and explainability about what goes into the algorithms and how these algorithms work. The model Nvidia, didn't look different from other autonomous cars, but it was unlike anything demonstrated by Google, Tesla, or General Motors. However, it showed the rising power of AI: the car didn't follow a single instruction provided by an engineer or programmer. Instead, it relied entirely on an algorithm that had taught itself to drive by watching a human do it.

Following is a fairly good podcast that gives some overview of Nvidia's self-driving cars.

- **How AI Will Revolutionize Driving — Danny Shapiro, NVIDIA (https://soundcloud.com/theaipodcast/how-ai-will-revolutionize-driving-danny-shapiro-nvidia)**

Having such a self-driving car seems impressive. However, the fact that no one indeed clearly knows how it made its driving decisions is a bit unsettling. Self-driving car sensors capture and collect a tremendous amount of data every single second. Such information about the environment has been feed straight into a huge network of artificial neural network model that process the data and then deliver the car operactions (control on the steering wheel, the brakes, and other systems) that required in react to the environment. From the demo, the result seems to match the responses you'd expect from a human driver. However, what if one day it did something unexpected—crashed into a tree, or sat at a green light?
Unfortunately, the system is often so complicated that it might be difficult to find out why. Even the engineers who designed it may struggle to isolate the reason for any single action.

Accountability, Responsibility and Governance

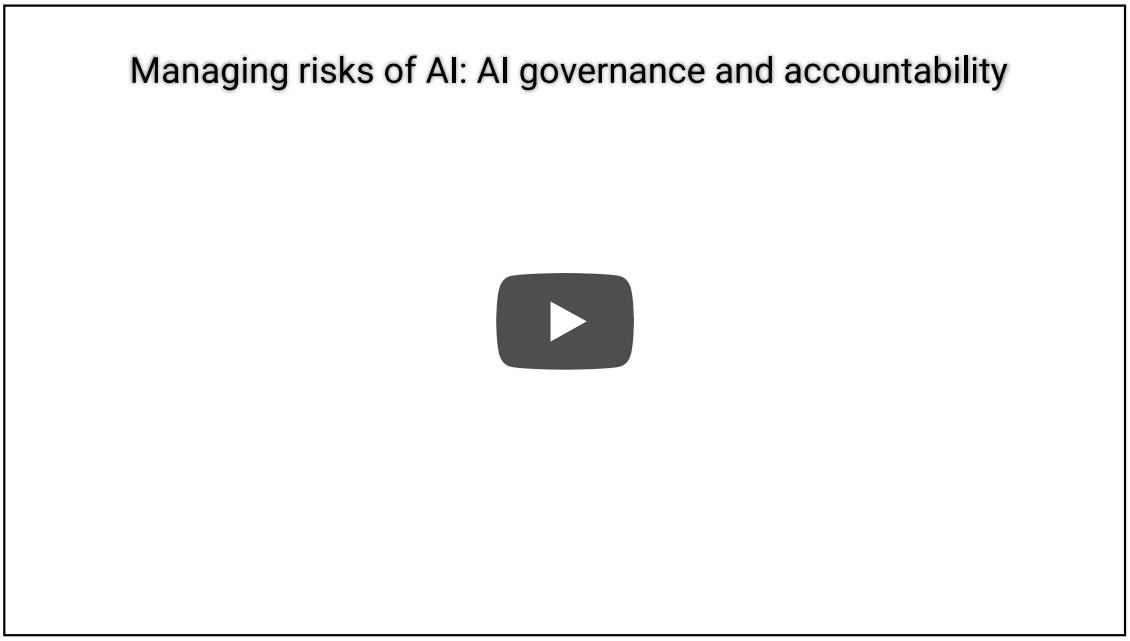
Developments in autonomy and machine learning are rapidly enabling AI systems to decide and act without direct human control. However, there is also a need for debate around who will be accountable for decisions made by AI.

"For there to be trust in AI, there needs to be a degree of accountability"
– Ms. Kate Marshall, KPMG Law partner

Greater autonomy necessarily comes with greater responsibility, although the notion of responsibility might be significantly different between when it applies to humans and when it applies to machines. How can governance functions keep up with fast-paced development in AI? Who should be held accountable?

Take self-driving cars as an example, as stated by **KPMG (https://home.kpmg/au/en/home.html)** 's technology partner Kate Marshall, the current state of rules surrounding the testing of driverless cars and autonomous vehicles are too reactionary itself, “with no overall framework or agreement around the approach to artificial intelligence being used”.

Responsible AI is more than the ticking of some of the ethical ‘boxes’ or the development of some add-on safety features in AI systems. It requires the participation and commitment of all relevant stakeholders and the active inclusion of all of society. This means training, regulation, and awareness, etc. Watch the following discussion between Cathy Cobey, Partner in EY Canada’s IT Risk & Assurance practice and Dr. Cindy Gordon, Founder and CEO, SalesChoice discuss the issues in relation to Managing the Risks of AI.



Human Right and Social Wellbeing

AI is everywhere and it’s here to stay. Most aspects of our lives are now touched by AI in one way or another, from operating home appliances, to deciding what movie to watch or flights to book online, to whether our job applications are successful, whether we receive a bank loan, and even what treatment a patient receives for cancer. All of these things – and many more coming – can now be determined largely automatically by complex AI systems. The enormous strides AI has made in the last few years are striking – and surely AI has the potential to make our lives better in many ways. However, AI can also be used to threaten human rights. For example, we have seen allegations of AI entrenching bias and discrimination in the United States (US) criminal justice system. So, all in all,

“The challenge now is to make sure everyone benefits from this technology”
– Peter Norvig, director of research at Google

Is there a need for a data science code of ethics? Why?

PRESCRIBED READING
Artificial Intelligence: Australia’s Ethics Framework (https://consult.industry.gov.au/strategic-policy/artificial-intelligence-ethics-framework/supporting_documents/ArtificialIntelligenceethicsframeworkdiscussionpaper.pdf) : Chapter 2 Existing frameworks, principles and guidelines on AI ethics

Responsible Data Science: **https://link.springer.com/article/10.1007/s12599-017-0487-z (https://link.springer.com/article/10.1007/s12599-017-0487-z)**

Workshop Discussion

- What impact is AI likely to have on Australian society in the near and longer-term?
- What are the ethical, trust and human rights challenges presented by AI?
- What are the AI Principles and standards that should be adopted for Australia?
- What does a governing body responsible for setting standards and guidelines for the ethical use of AI that would inform self-regulation and guide government regulation look like?