Data Ethics, AI and Responsible Innovation

Nick Hood

03 November 2020

Table of Contents

[Introduction 1](#_Toc55290578)

[Course overview 1](#_Toc55290579)

[Week 1: Ethics and Law 2](#_Toc55290580)

[Ethical ‘Issues’ 2](#_Toc55290581)

[Introduction to Ethics 3](#_Toc55290582)

[Appendices 4](#_Toc55290583)

[References 4](#_Toc55290584)

# Introduction

These are my notes from participation in the Edinburgh Data Ethics MOOC, running in November and December 2020. Any errors here are my own responsibility.

Nick Hood  
[@cullaloe](https://twitter.com/cullaloe)

This document was last updated on 03 November 2020 at 10:08.

## Course overview

Week 1 Ethics and Law

What are ethical values? Can we rely on legal regulation? What are the most pressing issues facing data driven industries?

Week 2 Crime and Justice

Should we use predictive policing and sentencing algorithms? How can biases sneak into such algorithms? How can we remove them?

Week 3 Home and City

What are the promises of smart homes and cities? How can they impact our privacy and freedom? How can we design them to protect those values?

Week 4 Money and Markets

Can future tech lead to a world without money? Can algorithms help us fairly distribute resources? How could we design a fair AI?

Week 5 Life and Health

Should we keep genetic databanks? Would you trust an AI doctor? What are the principles of responsible research and innovation?

# Week 1: Ethics and Law

This starts with some housekeeping, introductions to the tutors and setting up a profile within the course. Dr. Ewa Luger gets us started by considering a “broad overview” of **ethics**, using a radio interview about the Cambridge Analytica (CA) scandal as a resource to get us thinking. This is followed by an introduction to ethics

## Ethical ‘Issues’

In the clip, problematised issues with the CA matter included:

* the way data was gathered from people
* the fact that the data was then passed on to another party
* that the data was now available for commercial use
* that the data was used for political purpose
* the FaceBook (FB) login on an app gave permissions for all of users’ FB data to be used and shared
* people aren’t outraged, or even bothered by this enough to take action

Discussion forum task

Think back over your professional life and identify one example of an ethical issue that you personally experienced. In the discussion forum, write a brief summary (around 50-100 words) describing this issue - try to explain the context, what happened, and why you felt it was an ethical issue. After you’ve done this, comment on two other posts.

I had to take a quick look at the forum first to gauge the level of ethical matters that were being shared. There seems to be a discernible difference between **moral**, **legal** and **ethical** matters in the revelations on the forum. My own example came from quite a selection of professional experience that ranged from legal but immoral, moral but illegal, unethical but moral and legal. I picked an example from a while back.

I was project manager in a business that made flight crew training simulators. These devices were multi-million-dollar complex machines that used specially designed and manufactured circuit boards mounted in standard racks to drive the various components of the device. I had an internal review one morning that failed badly when the equipment that had been working well the night before suddenly failed to function. It was extremely embarrassing for my team. We discovered later that day that in the night, another project manager had switched a key circuit board for a faulty one from his own machine.

The other examples I commented on or read seemed to me to be not so much ethical issues as difficult choices or situations. This is causing me to wonder if I have a secure understanding of the term, “ethical”. Hopefully the next section will clear that up for me.

## Introduction to Ethics

Issues arising from the 2018 Cambridge Analaytica scandal are identified as:

* the value derived from data is not evenly distributed
* power rests in the hands of very few
* the potential for harm and inequality is high
* the potential impact of bad actors

These are, in my view, issues of power arising from wealth inequity and not particularly related to anything technological. The difficulty is that technology amplifies the effects (as it amplifies inequities in a classroom – this much we have learned in moving to “Digital First” pedagogies resulting from the COVID measures).

The basic principles of ethics are described as, “how we should address these issues, whilst minimising harm, ensuring morally good outcomes and fairness, and protecting human autonomy.”

### Data at the heart of it

For the purposes of the course, this definition of ethics is set in the context of *data-driven innovation* which makes use of large amounts of data to train algorithms that make predictions or provide insights for decision making. Machines are increasingly now involved in the decision making too, and in taking actions. These machines are described as *intelligent*.

### Visibility of rationale

These machines include, of course, neural networks and we know from the way that these pattern-matchers work, that other than the input and output layers, it is not possible to determine the path of the rationale – the specific choices ot selections being made in the intermediate steps – that underlies the outputs.

### Ambient intelligent systems

Interactional moments are perceived to provide friction in the operation of systems, therefore designers are removing these for a smoother experience – think of entering your password every time you wanted to look at your phone, now replaced with fingerprint or face recognition. The result is that such interactions become invisible to the user: it is the environment we interact with in a natural way. Unnaturally, *the environment is reading us to discern intent*.

### Human –> AI interaction

This is problematised in a video clip, suggesting that this interaction is not yet defined. I rather think that Turing ([1950](#ref-TURING1950)) had a clear enough view of what it looks like.

# Appendices

# References

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *LIX*(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>