DSRL – Cluster 5
Ethics – Part I

Noelle-Ann Park

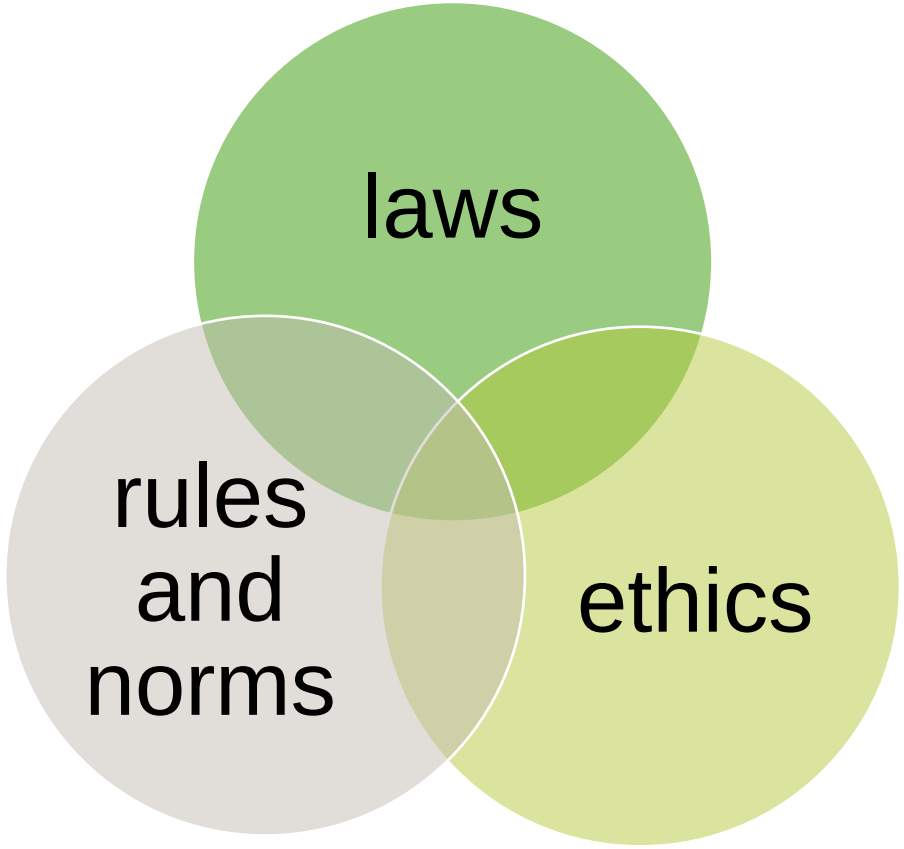
Objectives module

- Students become familiar with ethics, and how to apply this to data science practice
- Students are aware of the trade-offs and values involved in data science.
- Students become familiar with Ethical theories and their fundamental concepts and principles and learn how to apply this.
- Disclaimer

Thou shall not kill…..

What are ethics?

- Laws: describe what **can** and **cannot** be done
  - e.g. laws of physics, laws of countries, international conventions
  - Regulates the actions of people
- Rules and norms: describe what **may** and **may not** be done
  - e.g. ten commandments.
  - Standards of good behaviour.
- Ethics: various tools for figuring out what **should** or **should not** be done, what is **good** or **bad**
  - e.g. Being punctual, should I give money to a homeless person, thank the bus driver when I leave the bus.
  - What is morally right or wrong / just or unjust



Normativity and Relativism

- Ethical reasoning is not relativist, but normative: it results in a judgement about what **should** be done

- Relativism is about judging things in relation to a particular local framework, e.g. FGM is acceptable in some countries but not in Europe

- N.B. Relativism involves a normative judgement
  - ‘what is right depends on your perspective’
  - ‘there is no right answer to this question’

Analysis and processing of data – Steps involved.

- *Data gathering, preparation and exploration,*
- *Data representation and transformation,*
- *Computing with data,*
- *Data modelling,*
- *Data visualization and presentation.*

TILBURG UNIVERSITY

Donoho, D. (2017). 50 Years of Data Science. *Journal of Computational and Graphical Statistics*, *26*(4), 745–766. https://doi.org/10.1080/10618600.2017.1384734.

Emerging ethical problems of data science

1. **Free will and autonomy**: to what extent are people consenting freely to having their data collected, used and resold? *(In the context of AI - is a threat to free will?)*

2. **Visibility**: can we opt out of having data about us shared and used?

3. **Bias and discrimination:** AI bias, pattern recognition or classification and clustering. Negative impact on data subjects.

4. **Prediction, profiling and categorisation**: markets facilitated by algorithms can amplify inequalities and unfair treatment
5. **Collective impacts in a system based on individual rights**: all data, even de-identified data, tell a story about human activity that may have impacts on people

6. **Experimentation – (Accountability & Responsibility)**: 'perpetual beta' and a/b testing make users experimental subjects, leading to a loosening of the rules around experimentation.

7. **Quantification of human life:** all models make certain assumptions that then may influence how people see themselves

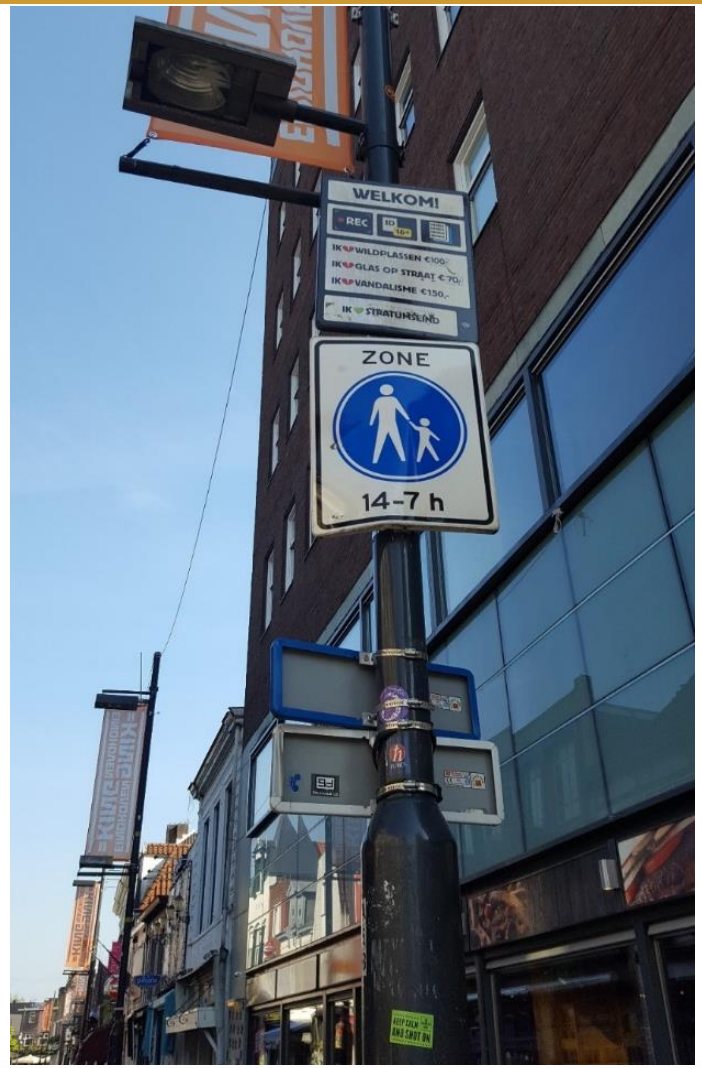
Tilburg University

1. Free will and autonomy - Informed Consent

- Freedom – act without constraint
- Autonomy
  - Independent desire.
  - often codified as ‘informed consent’
- Spotify: "We may collect information stored on your mobile device, such as contacts, photos, or media files.“

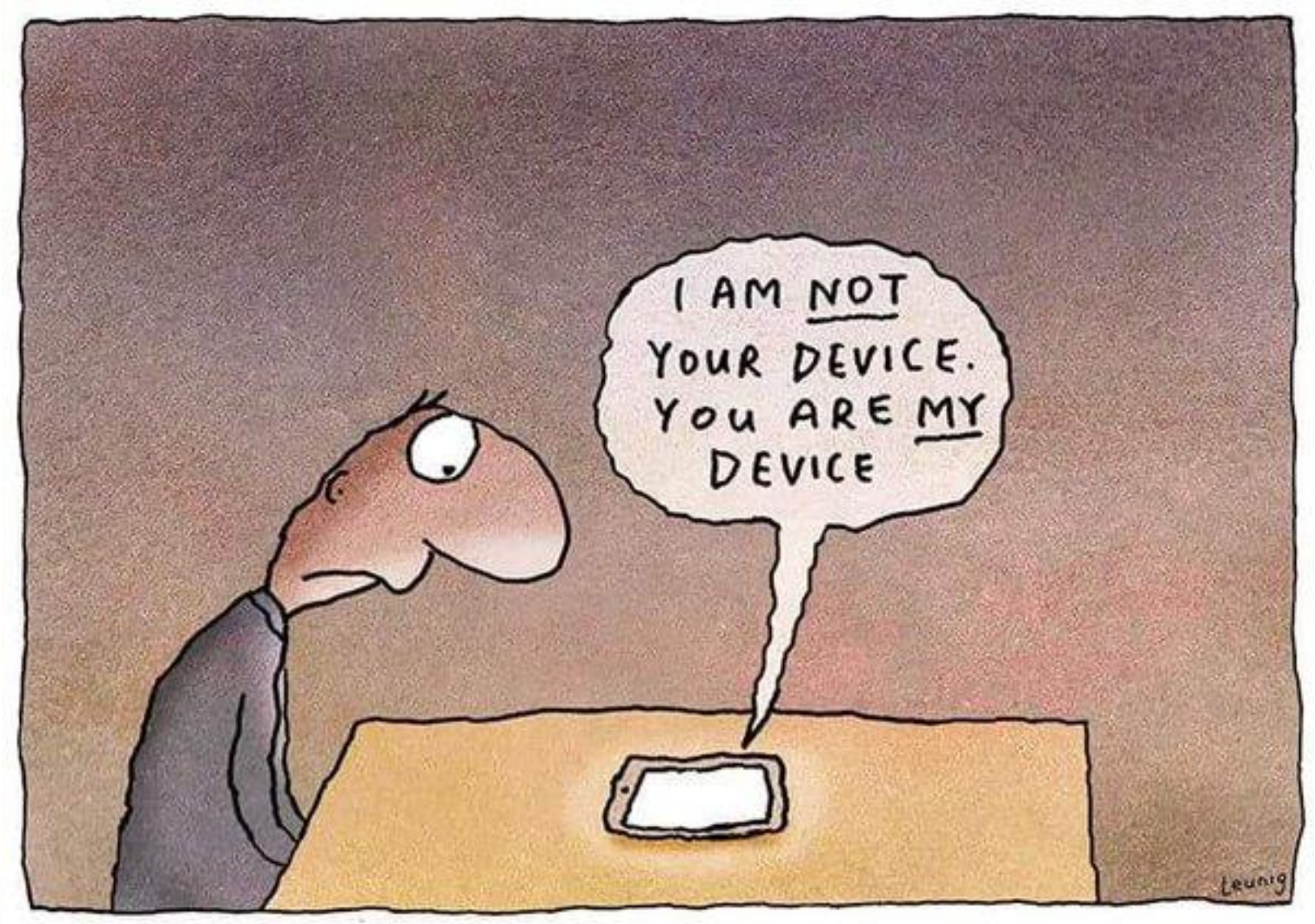
- Facebook: retains right to use your data for advertisements (including after account deletion), psychological studies that attempt to depress your mood, predicting and intervening when someone seems to be at risk of suicidal ideation and behaviour (although not in Europe)

Informed Consent



TILBURG UNIVERSITY

2. Visibility - Data Ownership



TILBURG UNIVERSITY

Opting out?

Facebook Is Tracking Me Even Though I'm Not on Facebook

By Daniel Kahn Gillmor, Senior Staff Technologist, ACLU Speech, Privacy, and Technology Project
APRIL 5, 2018 | 6:00 PM

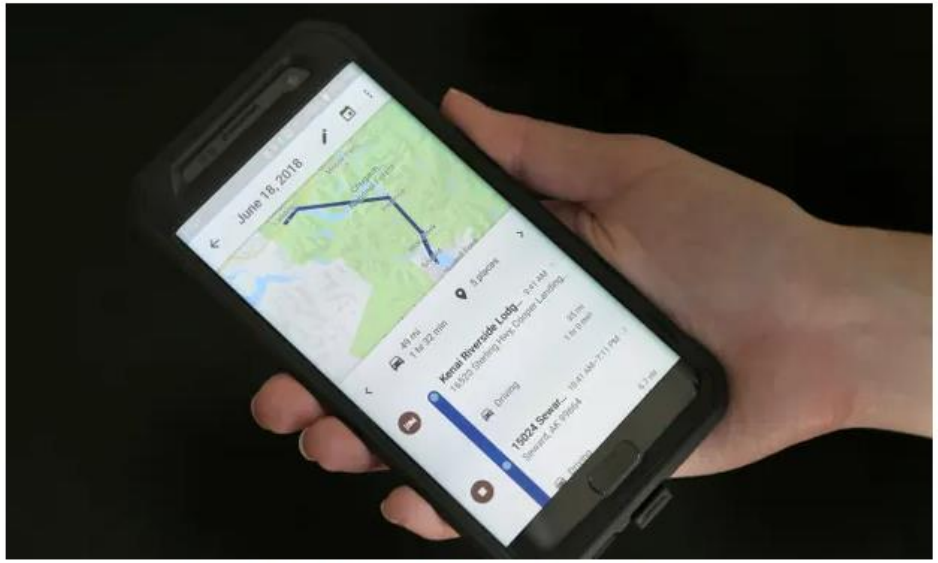
TAGS: Internet Privacy, Privacy & Technology, Consumer Privacy

I don't use Facebook. I'm not technophobic – I'm a geek. I've been using email since the early 1990s, I have accounts on hundreds of services around the net, and I do software development and internet protocol design both for work and for fun. I

https://www.aclu.org/blog/privacy-technology/internet-privacy/facebook-tracking-me-even-though-im-not-facebook

Google records your location even when you tell it not to

Some services on Android and iPhone automatically stores your movements even after you pause the 'location history' setting



A mobile phone displays a user's travels. Photograph: Seth Wenig/AP

Google wants to know where you go so badly that it records your movements even when you explicitly tell it not to.

Google records your location even when you tell it not to | Google | The Guardian

TILBURG UNIVERSITY

Privacy & security

‘Security- is the protection of information

- Privacy – right of individuals to keep their private information and activity private.
  - Privacy as control –
    - Individuals authority over their personal information –
    - choice to manage and limit access in collection of their own data give or not give consent.
    - People have control over their own privacy.
  - Are we really in control – TOS do we read Privacy policies.
  - Privacy as contextual integrity –
    - not a universal concept that is the same in every context.
    - Can be deeply tied in the context in which information is shared.
    - Depends on the norms an exceptions . Not a one size fits all.

- *“But even if we accept that clicking through a "Terms of Service" that no one reads can actually constitute true consent, even if we ignore the fact that these terms are **overwhelmingly one-sided** and **non-negotiable,** and even if we accept that it's meaningful for people to give consent when sharing data about other people who may have also opted in — what is the recourse for someone who has not opted into these systems at all?”*

The crowd makes the individual visible

‘Anonymity and informed consent emerged as panaceas because they presented ways to ‘have it all’; they would open the data floodgates while ensuring that no one was unexpectedly swept up or away by the deluge.’

But…

“multiple attributes can be inferred globally when as few as 20% of the users reveal their attribute information.” (Mislove et al. 2014)

‘As such, the value of any particular individual’s withheld consent diminishes incrementally the closer the dataset of those who granted consent approaches representativeness.’ (Barocas and Nissenbaum, 2014)

Data understanding & Anonymisation:

- Data collection followed by –
  - Data reading
  - And Data storing

“In statistics as well as machine learning toolboxes, correlations and distributions of feature values, diverse visualisations of the data, looking for typical or impossible feature values are the next steps to become acquainted with the data”

- *(Kietz, J.-U., & Morik, K. (1994)*

- Then how does Anonymization or other means of guaranteeing privacy impact how data is understood / processed.

Transparency and explainability

- What is transparency?
  - Disclosure of information
  - Openness and accessibility of Information
  - Example AI -inner working of the system must be clear to the user and understandable.
  - Operation should be clear and visible.
  - How and why certain decision are made
- But why Explainability ??
  - Targeted on taking the complex language - into an understandable language without leaving out important information that could affect the **Choice** of users.

3. Bias & Discrimination

- Bias –
  - Inclination or Prejudice
  - Against a person or group of people
  - Considered unfair
  - Bias in AI - *traditionally referring to the assumptions made by a specific model (Mitchell, 1997).*
- Discrimination:
  - The adverse effects that stems from bias.
  - Algorithmic discrimination –
    - Discrimination on the basis of specific criteria
    - Discriminative decision procedures

Examples:

TILBURG UNIVERSITY

Discrimination

- Positive discrimination –
  - Favours an individual or group of individuals
  - Treatment which is different and affects them positively – to their benefit.
    - TU/e Case study - In June 2020, Eindhoven University of Technology reserved all vacant positions exclusively for women for a number of years. The Eindhoven University of Technology revised the program subsequently.
- Negative Discrimination - explain with examples
  - Does not favour an individual or group of individuals.
  - Discrimination on the basis of ethnicity, religious gender.
    - SyRI decission - Dutch childcare benefits scandal. Case where false allegations of fraud made against parents. Partly based on ML model used the national benefits authorities that categorized those with **“dual nationality”** as a risk characteristic.

Algorithmic ‘mistakes’ are based on historical design decisions

- Algorithmic ‘mistakes’ are based on historical design decisions
- First film, then digital cameras were designed to distinguish **white skin from background**. Other colours of skin became ‘different’ from ‘normal’.
- Google Photos App did not recognise dark skin as human- Google apologises for Photos app's racist blunder - BBC News

TILBURG UNIVERSITY

Bias in data sets and implications on Public health.

- Sensitivity of Infrared Sensor Faucet on Different Skin Colours.
- Results suggested that there are biases between different skin-toned individuals in terms of the IR sensor faucet.
- Different reaction times between different skin shades with the faucet.
- Implications:
  - reduce proper hand hygiene practices.
  - Persons with skin shades induce a more prolonged reaction time more likely to avoid hand washing, making them more prone to illnesses.

Figure 1: Samples of hand-shaped Fitzpatrick Scale

Figure 3: SLOAN EAF-150 Faucet

TILBURG UNIVERSITY

Barassi, V. and Patra, R., 2022. AI Errors in Health?. *Morals & Machines*, *2*(1), pp.34-43.

4. Prediction, profiling, categorisation

Redlining: automatically disqualifying areas from service provision based on algorithmic decisionmaking about income/debt

Reverse redlining: automatically including areas based on the same process, for exploitative service provision (gambling, payday loans)

5&6. Experimentation

- Ubiquitous connectivity, cloud storage options and increased processing power lead to agile service provision and ‘perpetual beta’
- Testing and change are continuous
- Users become experimental subjects
- There are no controls on data science experimentation by corporations

- Ethics of experimentation on human subjects?
  - Straatumseind in Eindhoven.

Facebook’s 2014 mood experiment

2014: Facebook involves 600,000 users in a psychological experiment based on their news feeds

No subjects were aware of taking part, beyond their initial user agreement

A university review board consulted said the user agreement was enough because it incorporated A/B testing

UK world sport football opinion culture business lifestyle fashion environment tech travel

home › tech

Facebook

Facebook reveals news feed experiment to control emotions

Protests over secret study involving 689,000 users in which friends' postings were moved to influence moods

Poll: Facebook's secret mood experiment: have you lost trust in the social network?

This article is 3 years old

62,319 850

Robert Booth

Monday 30 June 2014 00.57 BST

TILBURG UNIVERSITY

Accountability & Responsibility

- Who is responsible? – about the task
  - Duty or the obligation assigned to a person
  - Their position at work
  - Does not mean you are fully responsible for the outcome
- Who is accountable? – outcome oriented
  - Answerability for the outcome / progress.

Example – AI - who is accountable?

Self driving car?

Real human that is accountable for the output of this systems?

7. Quantification of human life

- All models contain assumptions about human behaviour and human beings
  - For instance which variables are important to track
  - Or how human behaviour is conceptualized (e.g. in terms of likes and dislikes and the aggregation of these)
- 10.000 steps
- 2000 (or 1800) calories a day; 5 a day
- Such quantification might do something (negative) to the way we view ourselves and the way we behave – e.g. exercise is about performance, rather than enjoyment
- Quantification of the social world; The Metric Society – FB friends, social network [Mau, 2019]
- Data is connected to the people.

What is the “right” thing to do?

- “Right” ?
  - What according to you is right?
    - ***Good or just***.

- ***For example, if you as a developer chooses which datasets to use in order to minimise harm.***

- What questions would you ask in this case in terms of what is the “right” thing to do?

What do we need to focus on Ethics in Data science?

- Acquiring data sometime without even having met the data subject.
- Data may not be identifiable – but can still be risky (data leaks)
- Accountability & Responsibility
- Trust
- Transparency

Linnet Taylor's Justice in Ethics Research

Visibility

Can you own your own visibility?

Can you withdraw when visibility and transparency are not beneficial and clear?

Engagement with Technology

Is there a choice to engage with freely?

Actual choice!

Non-Discrimination

Is this something that users / people must identify themselves?

Should those responsible for the technology be accountable here?

Tilburg University

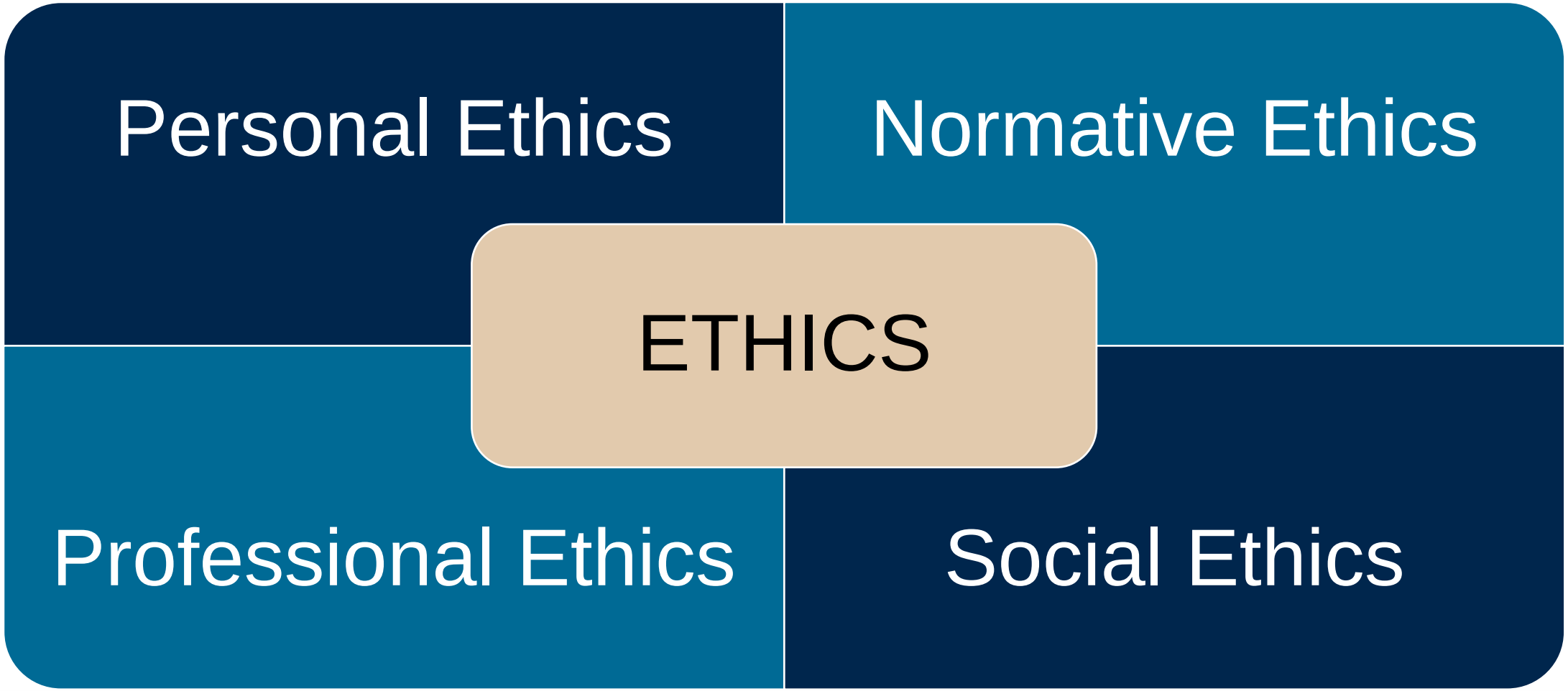
Any questions so far?

Tilburg University

Part II Ethical Theories

Brief re-cap of Part I

- Became familiar with ethics, and how to apply this to data science practice:
  - Free will and autonomy
  - Visibility
  - Bias and discrimination
  - Prediction, profiling and categorisation:
  - Experimentation & Collective impacts in a system based on individual rights:
  - Quantification of human life.
- The trade-offs and values involved in data science.


Personal Ethics
Normative Ethics
ETHICS
Professional Ethics
Social Ethics

Content for Today

- Ethical theories and their fundamental concepts and principles:
  - Deontology
  - Consequentialism
  - Virtue Ethics

Theory	Basis
Deontology	The act in itself is good
Consequentialism	What will happen if I do this? (cost / benefit)
Virtue ethics	What would a good person do

Deontology

- Deon – means duty or obligation – Immanuel Kant
- According to Deontology –
  - Morality of a person stems of the human action
  - And if that action fulfills a duty or not.
  - Some actions in themselves are either ***wrong or right .***
  - It is therefore our ***duty*** to perform these ***right actions.***
  - Examples:
    - Telling the truth under oath –
    - So even if there are possible negative / unpleasant consequences – the act of truth telling inself is moral.
    - And avoid telling lies.
    - So if the action in itself is right – we have a moral duty to do the right thing.

- According to Kant –
  - Human being have the ability to think and reason
  - The Focus is on duties, defined by right and wrong.
  - And not on the consequences.
- Right & Wrong – are distinct from Good & Bad
  - Right actions – what we ought to do or are morally required to do
  - Wrong actions – what we ought not to do or are forbidden from doing
- The Right Action – ultimately the rational action
- Kant believes that ethics consists of commands about what we ought to do.
- “Imperative” - a command or order.

Examples

- Hypothetical imperative is a contingent command:
  - *“If you want/need ABC, then you ought to do XYZ.”*

- Moral imperative not contingent on wants, desires, or needs. (categorical imperative)
  - “Do ABC,” or
  - “You ought to do ABC.”

Criticism of Deontology

- The assumption that universalizing a maxim permits an action.
  - “I will lie whenever it’s convenient to get what I want,”
- So people would constantly lie in order to achieve an end.
- Could result in the concepts of “lie” and “truth” becoming meaningless.

Consequentialism

- Consequentialism – depends only on the cosequenes.
  - Utilitarianism –
    - act is morally right if and only if that act maximizes the good,
    - if the total amount of good for all minus the total amount of bad for all is greater
      - » Jeremy Bentham (1789), John Stuart Mill (1861), and Henry Sidgwick (1907)
  - Hedonism
    - pleasure is the only intrinsic good and that pain is the only intrinsic bad.

an act is right if and only if it causes "the greatest happiness for the greatest number."

Consequentialism and its various combinations:

- Consequentialism - act is morally right depends only on consequences.
- Actual Consequentialism - act is morally right depends only on the actual consequences (not for unforeseeable / possible outcomes)
- Maximizing Consequentialism - depends only on which consequences are best (not just satisfactory or an improvement)
- Hedonism - the consequences depends only on the pleasures and pains in the consequences (as opposed to supposed goods, such as freedom, knowledge etc.)

Criticism of Consequentialism

- Why would Consequentialism be problematic in data science?
  - Cost benefit?
  - Benefit to the company Vs. Benefit for the people
  - There is uncertainty in the “consequences” of our actions – predicting the future.

- Regardless of what the actual consequences may be, the morally correct thing for an individual to do is whatever they believe at that point will result in the best consequences.
    - How do you measure what outcome is the best
    - It could lead to people justifying wrong-doings.

Examples

1. You are developing an algorithm for a Bank for predicting who is in most need of loan facilities. You use income, place of residence and number of person in a household as a base profile.

2. A university asks you to help develop a system they can use to help students in need of tutoring facilities. You use previous academic scores and language proficiency as a testing variable.

3. You are developing an algorithm to detect false/incorrect information. This leads to some posts of users being blocked/removed. The removal/deletion of posts is based on accurate / factual information provided to you by a Government agency.

Virtue Ethics

- **Primary for Virtue Ethics –**
  - The good is the fundamental category
  - The good person = the person of good character
  - He who possesses moral virtue
- Aristotle – earliest exponents of the Virtue Theory
  - Distinction between theory and practice:
  - Knowing what is right and wrong on one hand BUT,
  - Actually doing the right and avoiding the wrong on the other hand.
  - Acquire virtue through practice

So what should Virtue Ethics look like?

- Focuses upon the character of moral agents
  - Rather than on the moral status of their actions or
  - the consequences of those actions.
  - The good internal qualities a person embodies – then is manifested into actions, practices and attitude.
  - Virtues include courage, kindness, honesty, justice, and temperance.

The principal virtues and vices

Sphere of feeling or action	Excess (vice)	Mean (virtue)	Deficiency (vice)
Fear and confidence	Rashness	Courage	Cowardice
Pleasures and pains	Self-indulgence	Temperance	Insensibility (rare)
Getting and spending (minor)	Prodigality	Liberality	Meanness
Getting and spending (major)	Tastelessness	Magnificence	Niggardliness
Honour and dishonour (major)	Vanity	Proper pride	Pusillanimity
Honour and dishonour (minor)	Ambition	Proper ambition	Lack of ambition
Anger	Irascibility	Good temper	Lack of spirit
Self-expression	Boastfulness	Truthfulness	Mock modesty
Conversation	Buffoonery	Wittiness	Boorishness
Disposition to others	Obsequiousness	Friendliness	Cantankerousness
Shame	Bashfulness	Modesty	Shamelessness
Indignation	Envy	Proper indignation	Spite

Criticism of Virtue Ethics:

- The focus – human happiness or flourishing
  - So? Is the job of ethics to tell us what we are already supposed to know?
  - The job of ethics is to give an account and justification of why certain behaviors or actions are either **good or bad**, virtuous or vicious, right or wrong.
  - Why Ethics ?
    - It is because we do not know, or are unsure about, what we should do or how we should live that we turn to ethics in the hope of finding answers.
    - This is what the standard modern theories of ethics, **utilitarianism and Kantianism**, attempt to do;

Examples for Virtue ethics

- *You receive and email for which you are not the intended recipient.*
- *The email contains some exam questions and solutions for an upcoming course exam.*
- *The course is particularity challenging and many classmates including you are not well prepared.*
- *What would you do?*

Example – conflict between Deontology & Consequentialism

- You work in an organization handling personal information (including Social Security numbers and driver's license numbers).
- You flag a data breach. The hackers want to be a paid a sum of money and in return they will delete the data and not release it to the public.
- What do you do –
  - Pay off the hackers to delete the data and keep the breach quiet?
  - Do nothing at all ?
  - Inform all your users about the data breach and risk their personal information being released?

Ethics with Mark White and Robert Arp – Batman Vs. the Joker

- Utilitarianism
  - Would kill the joker
  - By taking one life – many are saved
- Deontology
  - Reject killing the joker
  - Killing is wrong
  - The act in itself is bad
- Virtue Ethics
  - Would look at the character of the person who wants to perform the action – Batman
  - What is his character - does he want to be the person who kills?

Any questions so far?

Tilburg University

End of Part II Ethical Theories