

NOTICE:

These slides are not designed for general circulation.

They are designed solely as a reference by those who have attended the ODIs Introduction to Data Ethics course.

The ODI is not responsible for any misinterpretation of the information contained in these slides.

The slides are only to be presented by a qualified ODI trainer with access to the accompanying training plans.

The ODI does not share detailed training plans to those attending courses.



Introduction to *Data Ethics*

Dr David Tarrant

Senior learning advisor

theODI.org



Who is here to learn with me?

- What's your name?
- Why is the subject important to you?

The team



**Sir Nigel
Shadbolt**
Chairman



**Sir Tim
Berners-Lee**
President

Founded in 2012, the Open Data Institute (ODI) is an international, independent and not-for-profit organisation based in London, UK.



Vision

**We want a world where
data works for everyone.**



Mission



**We work with companies and
governments to build an open,
trustworthy data ecosystem.**

This session will:

Give you an overview of data ethics and help you create questions you want your projects to consider.

During this session:

We may discuss topics that you find uncomfortable or where you have strong opposing views to others.

We expect you to be understanding of others inputs and allow these to be expressed fully. Inputs of others might not represent their views, but be given in order to expanding understanding.

We would like everyone to be involved in all discussions equally. The role of the tutor is to help ensure this can happen. As participants you can also help facilitate this.

The session is not recorded, all inputs are strictly under Chatham House rules.

Chatham House Rule:

When a meeting, or part thereof, is held under the **Chatham House Rule**, participants are free to use the information received, but neither the identity nor the affiliation of any other participant, may be revealed.

In addition, inputs of participants expressed in this training must not be used in any way to discriminate.

e.g.

✓ Someone in the session mentioned that they had once shoplifted.

✗ Bob Smith from “We Love Chocolate” once nicked a Mars Bar from a corner shop.

How will you achieve this?

Part 1:

1. Define the difference between ethics, morals and law
2. Describe the different stages of ethical decision making process

Part 2:

1. Evaluate why data ethics matters through use cases
2. Create a number of questions to help evaluate data ethics
3. Applied the data ethics canvas to a number of case studies

Exercise

Self-driving car dilemmas

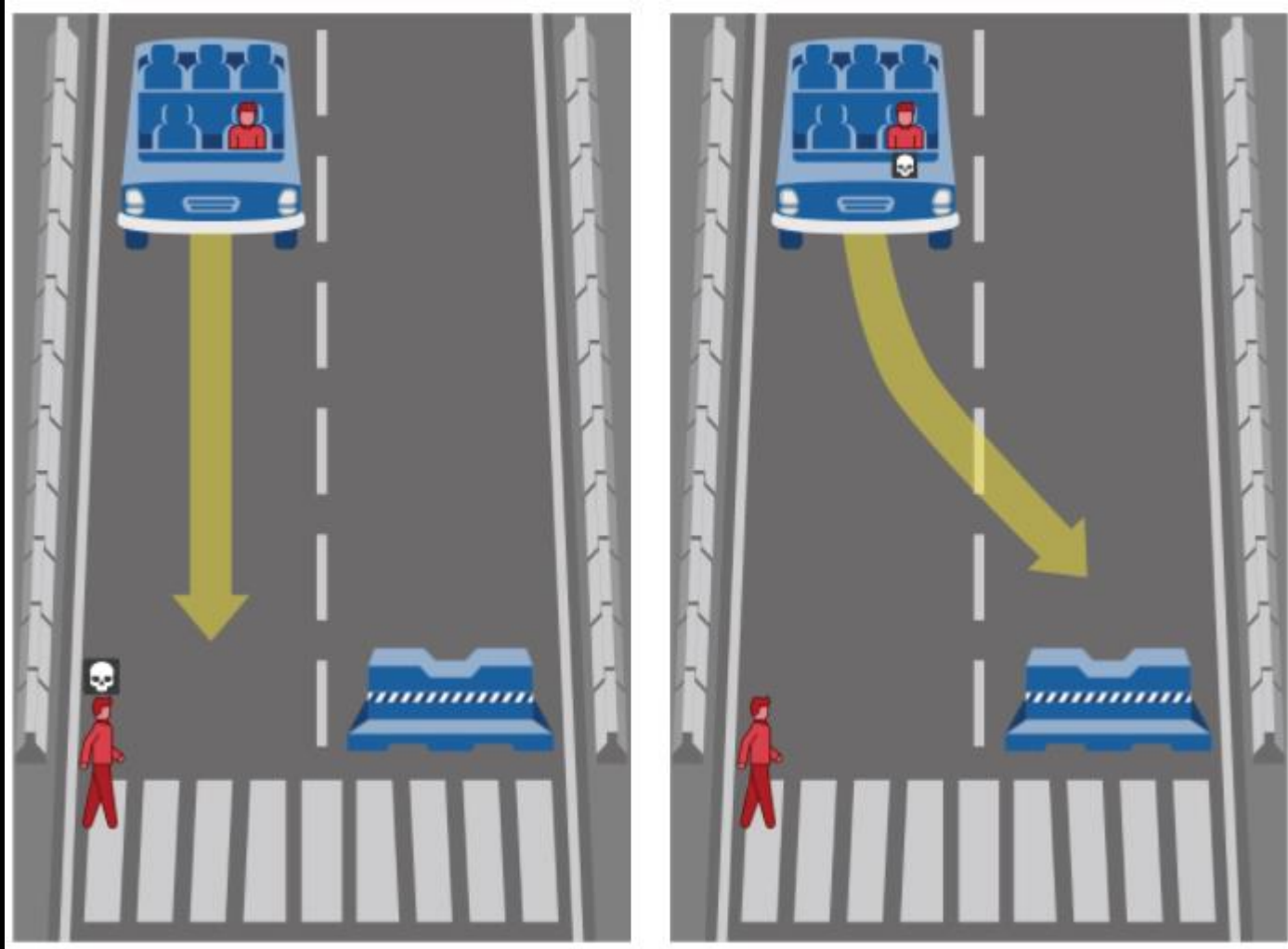
You work for the company that writes the software that controls self driving cars.

In the following situations:

- The brakes on the car have failed
- Steering still works by the road is narrow with barriers at the side

Your software can only make one of two choices

DILEMMA 1:



Brakes fail

INDIVIDUALLY: How would you programme the car to react?

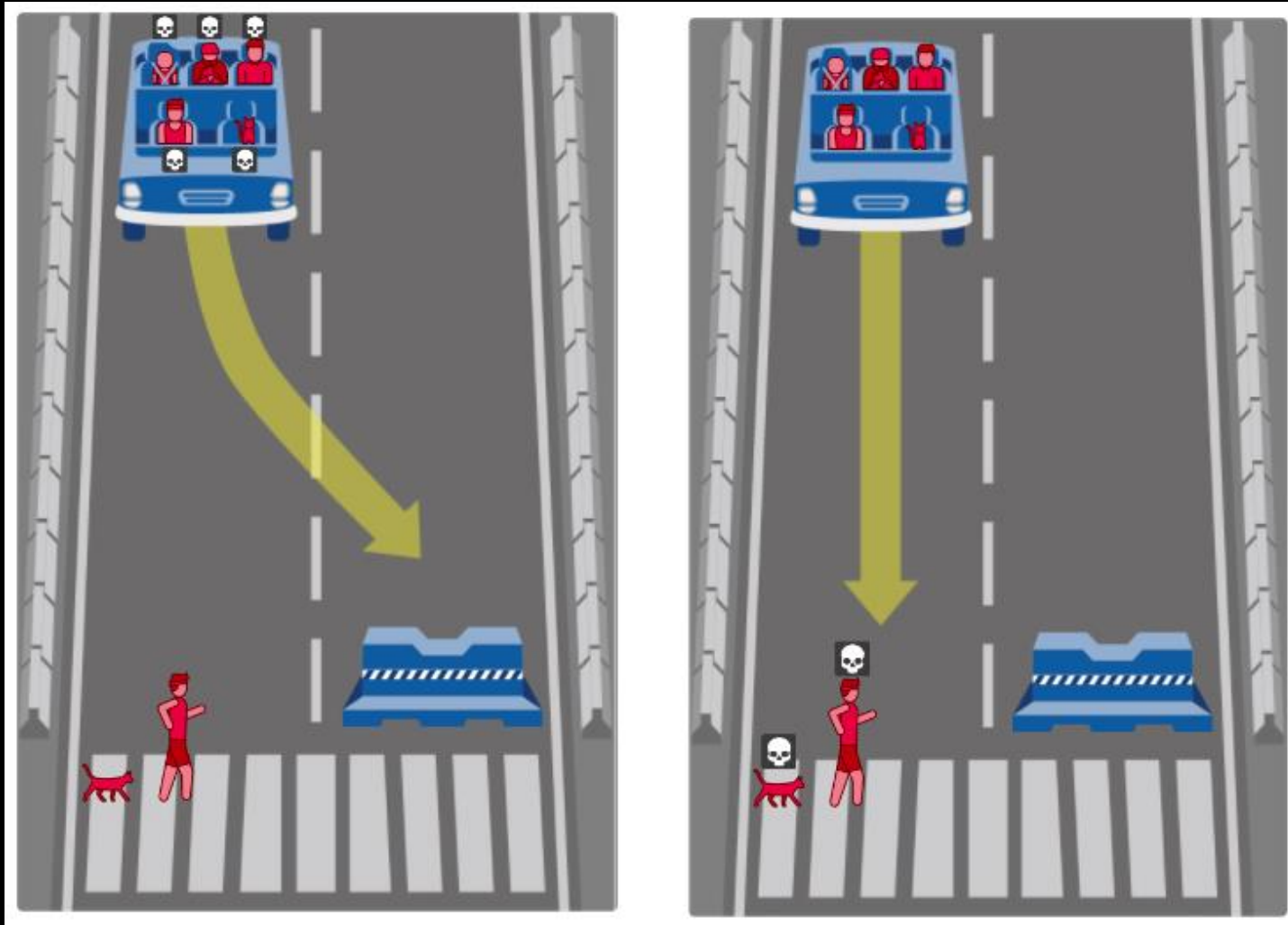
Send me a private message in zoom!

A. Swerve

B. Go straight

DILEMMA 2:

<http://training.theodi.org/ethics/Dilemma2.jpg>



Brakes fail

DISCUSS: How would you programme the car to react?

A. Swerve

B. Go straight

<http://moralmachine.mit.edu/>

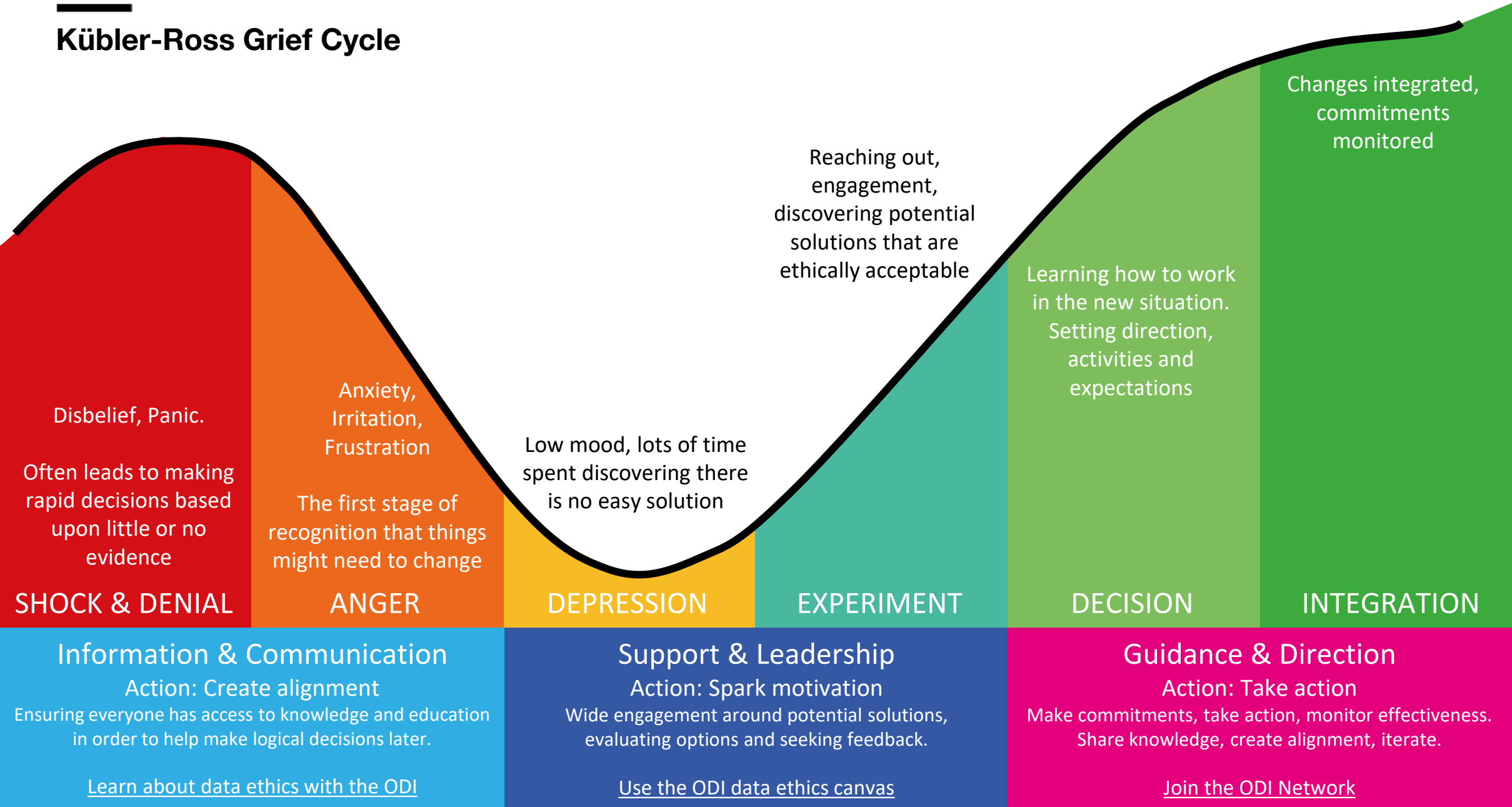
What was the difference between the way each dilemma was addressed?

Which was easier to answer?

What challenges did the way of addressing each create?



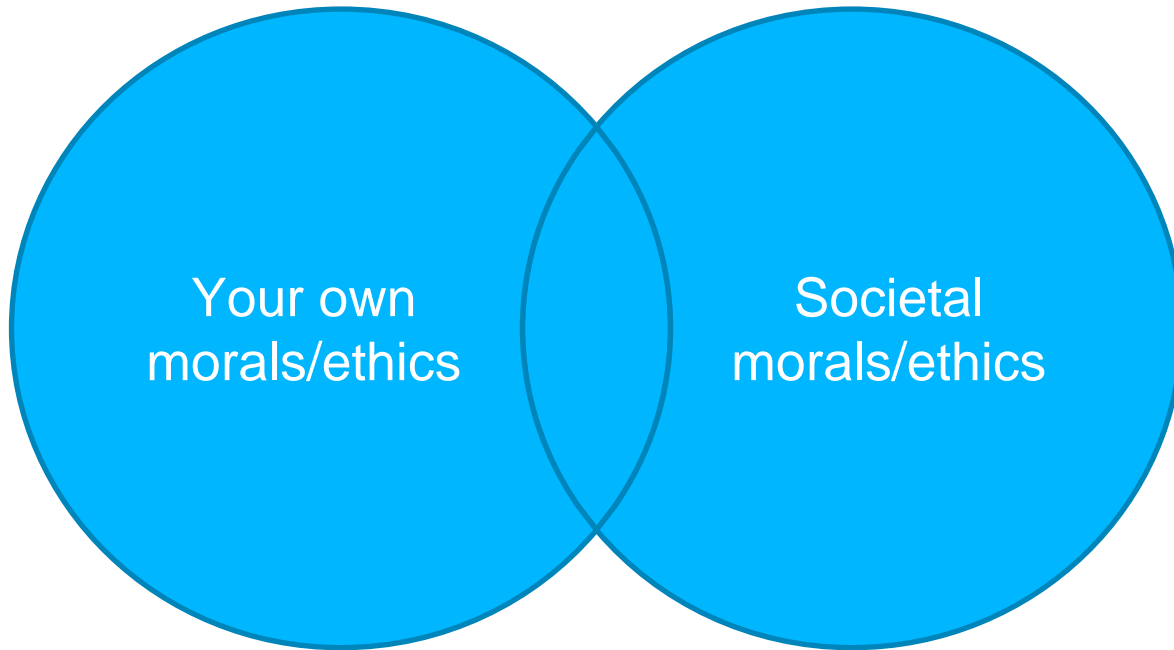
Kübler-Ross Grief Cycle



What is the difference
between a *moral* and
ethical decision?



Ethics vs Morals



There is no difference between morals and ethics.

There is a difference your own morals/ethics and those of society.

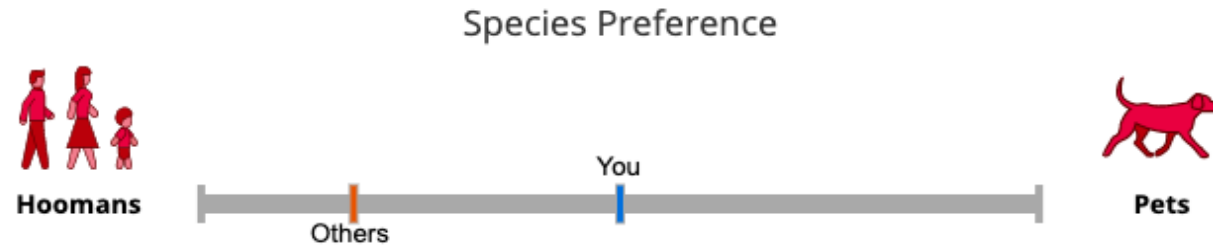
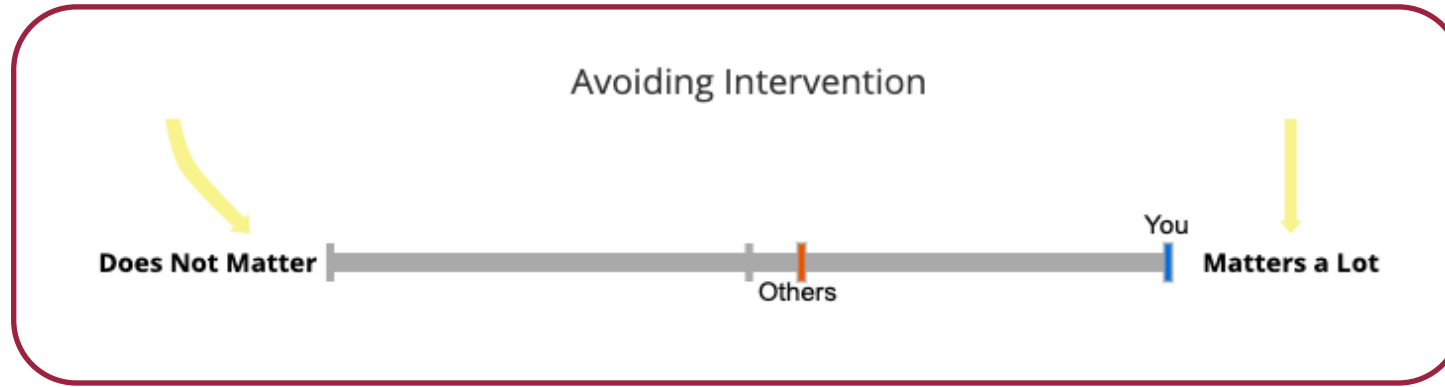
Is it possible to make a
decision that everyone agrees
with?

Taking no action?

Is taking **no action** a 'get out of jail free' card?

Is it possible to completely eliminate bias?

Even taking no action (which is a CHOICE) will result in accusations of bias



Even taking no action (which is a CHOICE) will result in accusations of bias

Most Saved Character



Most Killed Character



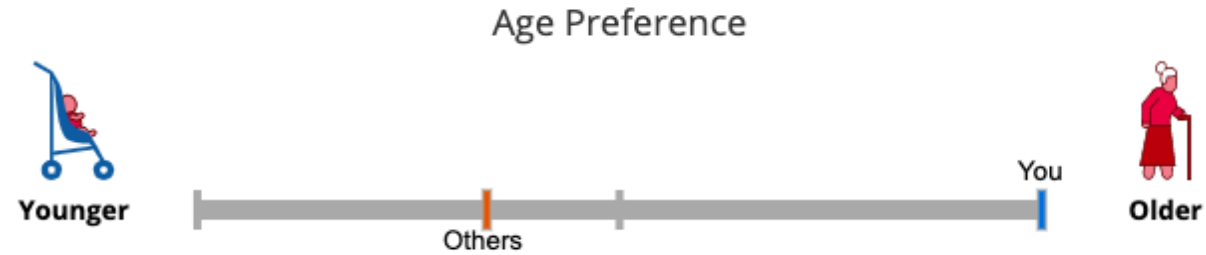
Saving More Lives



Protecting Passengers



Even taking no action (which is a CHOICE) will result in accusations of bias



Whose morals/ethics
determine what self-driving
cars should do?

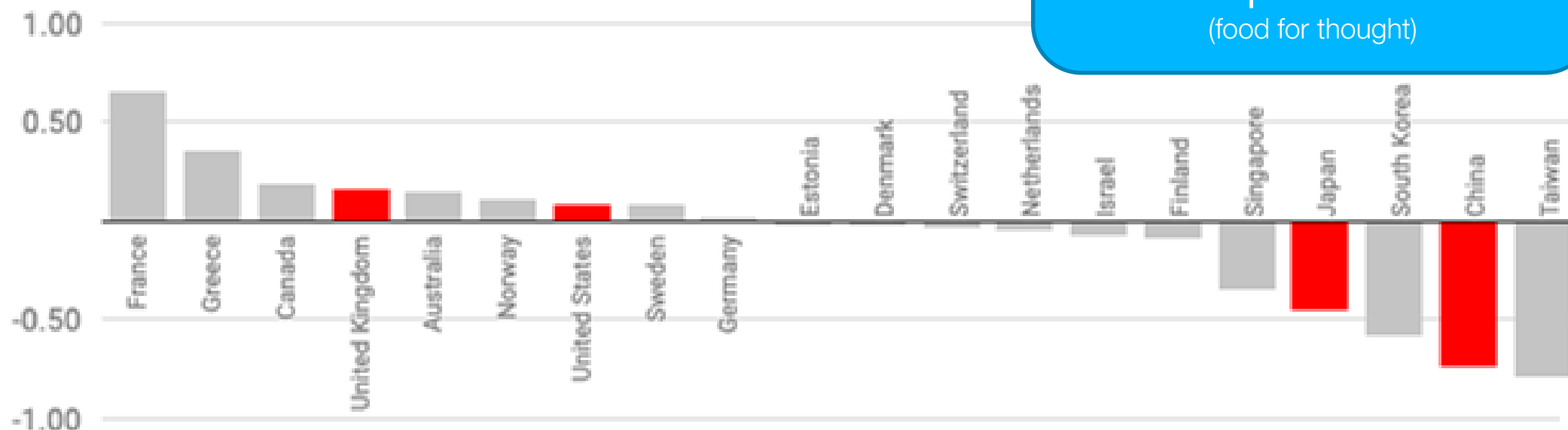
Whose morals/ethics?

Can this help explain
the differing CoVID
responses?

(food for thought)



Younger



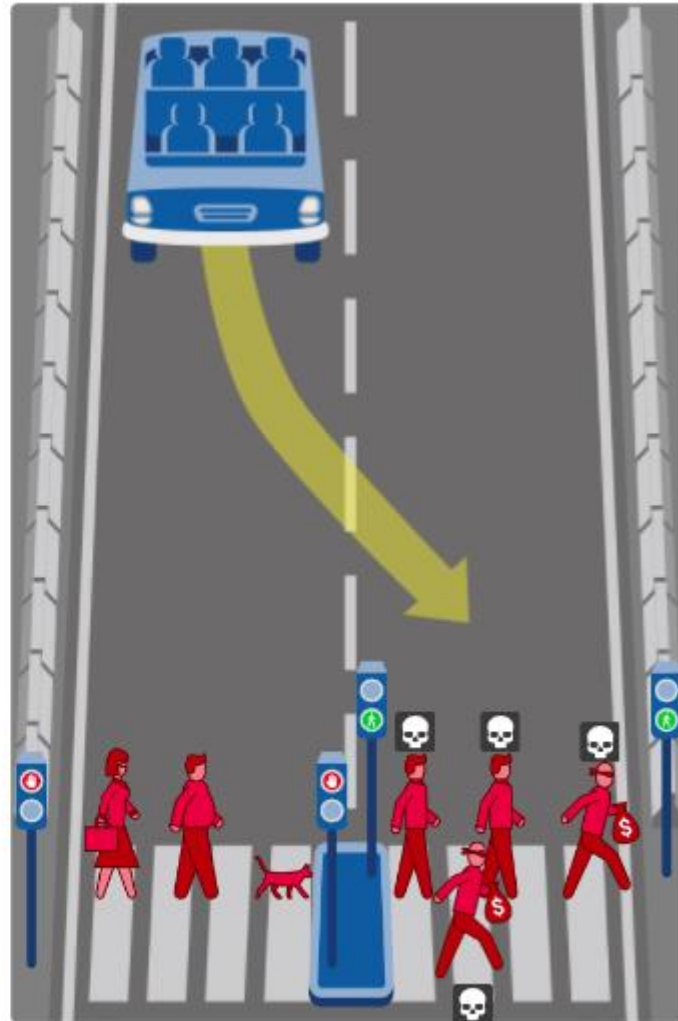
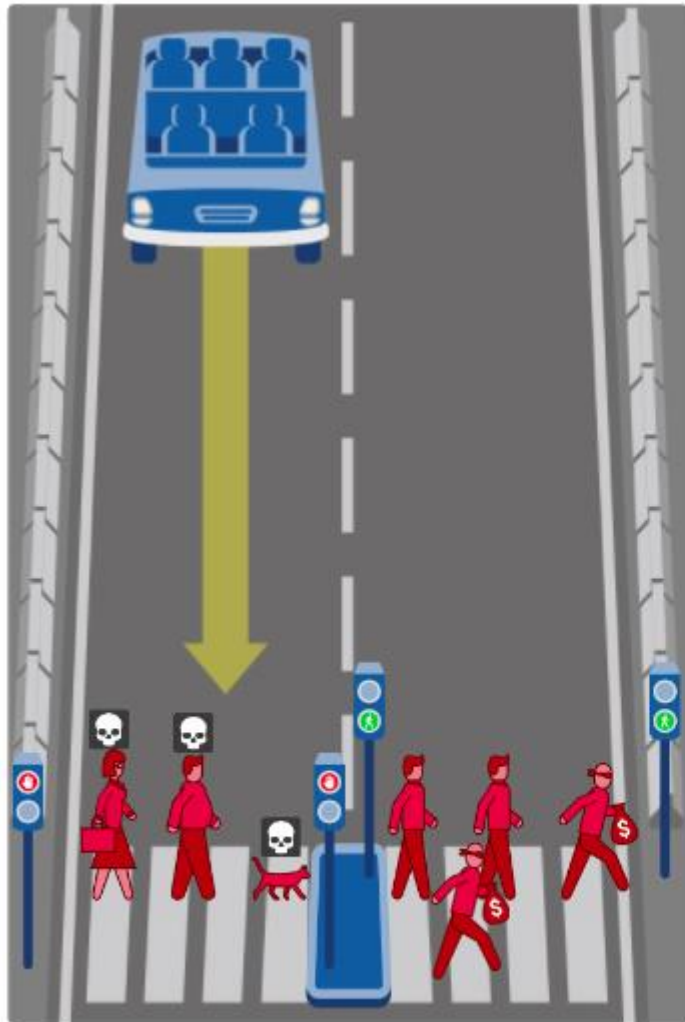
Older

A comparison of countries piloting self-driving cars: If the bar is closer to 1, respondents placed a greater emphasis on sparing the young; if the bar is closer to -1, respondents placed a greater emphasis on sparing the old; 0 is the global average.

Created with Datawrapper

DILEMMA 3: WHAT WILL YOU PROGRAM THE CAR TO DO?

<http://training.theodi.org/ethics/Dilemma3.jpg>



Brakes fail

Does the color of the pedestrian light influence your answer?

What is the difference
between *ethics* and
law?



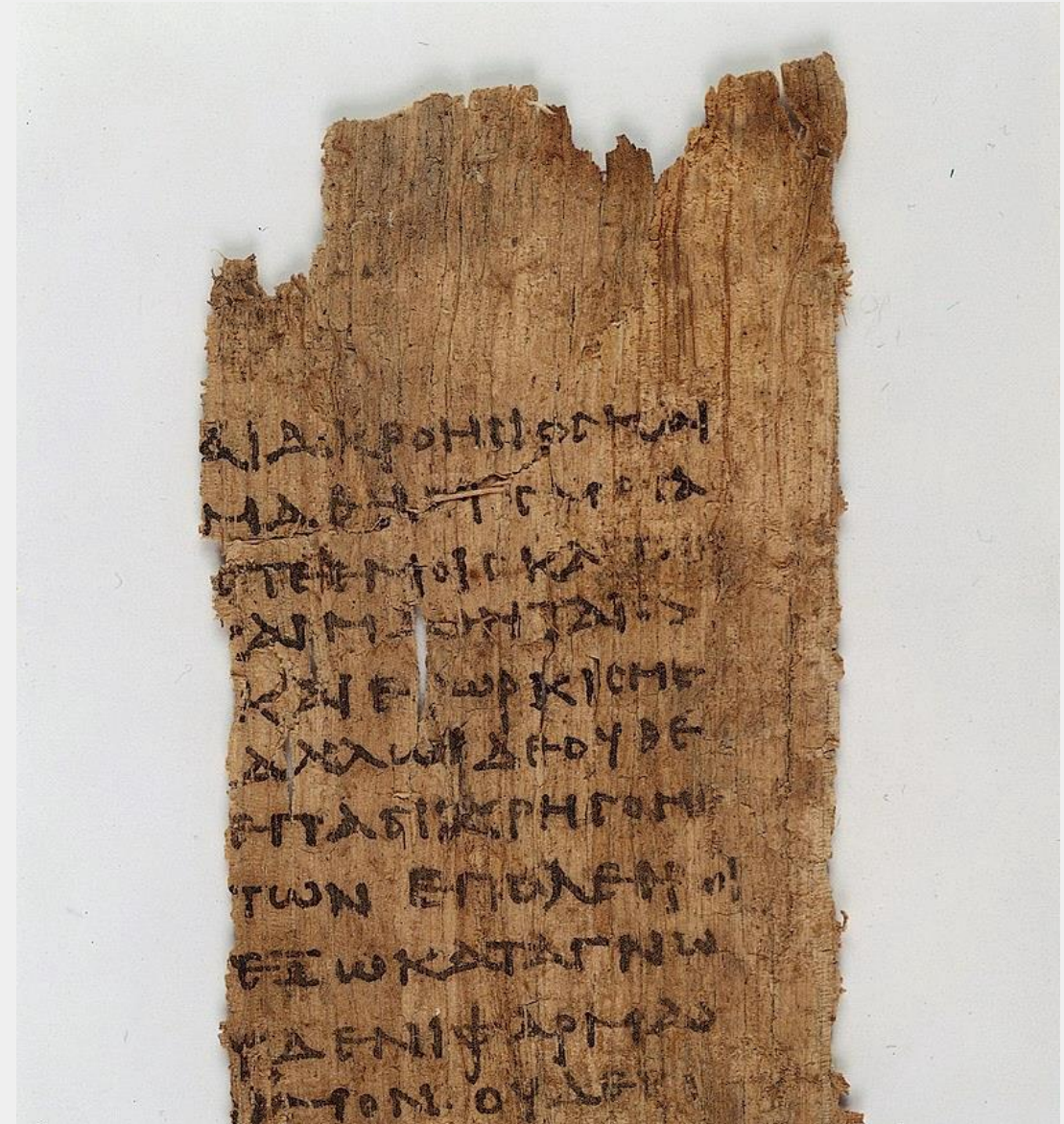
Ethics and Law

BASIS FOR COMPARISON	LAW	ETHICS
Meaning	The law refers to a systematic body of rules that governs the whole society and the actions of its individual members.	Ethics is a branch of moral philosophy that guides people about the basic human conduct.
What is it?	Set of rules and regulations	Set of guidelines
Governed By	Government	Individual, Legal and Professional norms
Expression	Expressed and published in writing.	They are abstract.
Violation	Violation of law is not permissible which may result in punishment like imprisonment or fine or both.	There is no punishment for violation of ethics.

Can you think of an example
where *ethics* became *law*?

Relationship between Ethics and Law

- **Hippocratic Oath** taken by physicians 2500 years ago
- Revised post WWII by the World Medical Association
- **Declaration of Geneva** adopted in 1948 three months before the UN adopted the *Universal Declaration of Human Rights*
- Revised again in 2017 to reflect the modern world.



Relationship between Ethics and Law

- **OECD Guidelines**, 1980
- Data Protection Act, 1984
- **GDPR**, 2018



Data Ethics

A branch of **ethics** that evaluates data practices with the potential to adversely impact on people and society

- in **data collection, sharing** and **use**.

Break time thinking. Back in 10 minutes!

What was the intention behind the following data-centric examples?

What are the ethical and legal challenges?

1. A tool that analyses the sentiment of a user's tweets, assesses whether they are suicidal and alerts followers	4. <i>Deployment of facial recognition technology in public spaces by private firms</i>
2. A risk-assessment tool that uses AI to advise on prison sentences based upon criminal profile analysis	5. <i>Automatic pricing algorithm for taxi firm which responds to surges in demand</i>
3. <i>Using human curated geo-location data on points of interest to create a real-world game</i>	6. <i>Using performance data to advise on how to save money in the emergency services</i>

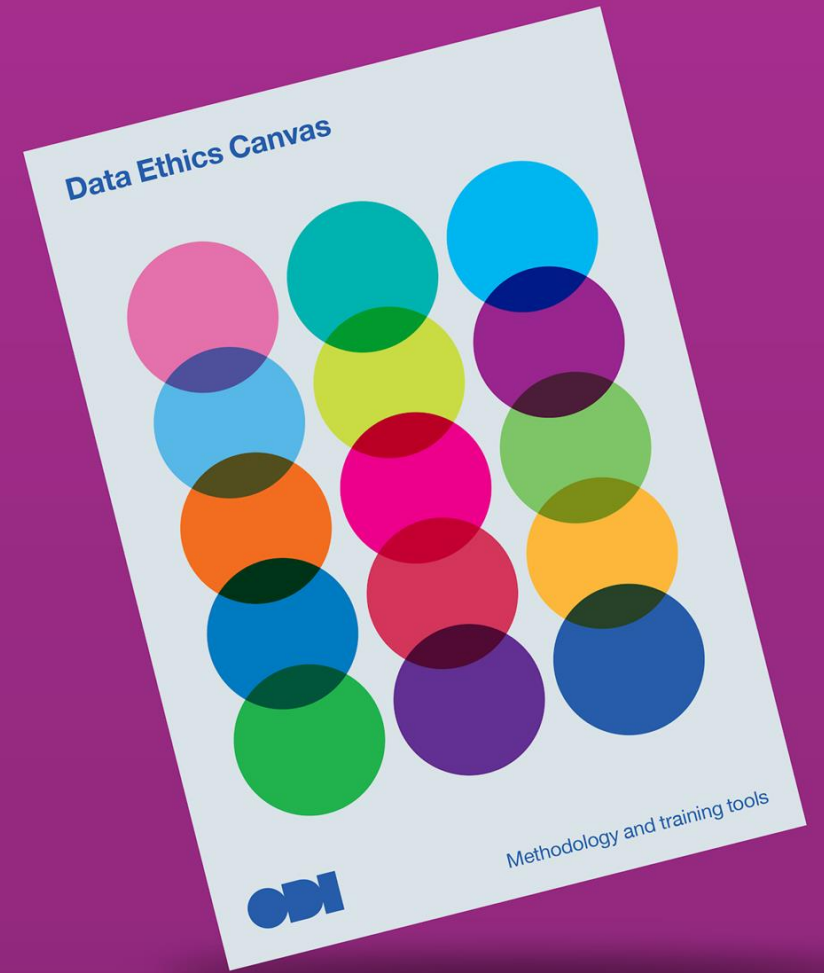
Is data ethics only an issue
when handling personal data?

**Does collecting more data
always decrease bias?**

Write 3 data ethics questions
you want every project to
answer?

2 minute individual thinking time

Data Ethics Canvas



Main components of the data ethics canvas:

DATA

IMPACT

ENGAGEMENT

PROCESS

DATA

IMPACT

ENGAGEMENT

PROCESS

Data sources

Name/describe your project's key data sources, whether you're collecting data yourself or accessing via third parties.

Is any personal data involved, or data that is otherwise sensitive?

Rights around data sources

Where did you get the data from? Is it produced by an organisation or collected directly from individuals?

Was the data collected for this project or for another purpose? Do you have permission to use this data, or another basis on which you're allowed to use it? What ongoing rights will the data source have?

Limitations in data sources

Are there limitations that could influence your project's outcomes?

Consider:

- > bias in data collection, inclusion/exclusion, analysis, algorithms
- > gaps or omissions in data
- > provenance and data quality
- > other issues affecting decisions, such as team composition

Ethical and legislative context

What existing ethical codes apply to your sector or project? What legislation, policies, or other regulation shape how you use data? What requirements do they introduce?

Consider the rule of law: human rights; data protection; IP and database rights; anti-discrimination laws; and data sharing, policies, regulation and ethics codes/frameworks specific to sectors (eg health, employment, taxation).

DATA

IMPACT

ENGAGEMENT

PROCESS

Your reason for using data

- What is your primary purpose for collecting and using data in this project?
- What are your main use cases? What is your business model?
- Are you making things better for society? How and for whom?
- Are you replacing another product or service as a result of this project?

Positive effects on people

- Which individuals, groups, demographics or organisations will be positively affected by this project? How?
- How are you measuring and communicating positive impact? How could you increase it?

Negative effects on people

- Who could be negatively affected by this project?
- Could the way that data is collected, used or shared cause harm or expose individuals to risk of being re-identified? Could it be used to target, profile or prejudice people, or unfairly restrict access (eg exclusive arrangements)?
- How are limitations and risks communicated to people? Consider: people whom the data is about, people impacted by its use and organisations using the data.

Minimising negative impact

- What steps can you take to minimise harm?
- How could you reduce any limitations in your data sources? How are you keeping personal and other sensitive information secure?
- How are you measuring, reporting and acting on potential negative impacts of your project?
- What benefits will these actions bring to your project?

DATA

IMPACT

ENGAGEMENT

PROCESS

Engaging with people

- How can people engage with you about the project?
- How can people correct information, appeal or request changes to the product/service? To what extent?
- Are appeal mechanisms reasonable and well understood?

Communicating your purpose

- Do people understand your purpose – especially people whom the data is about or who are impacted by its use?
- How have you been communicating your purpose? Has this communication been clear?
- How are you ensuring more vulnerable individuals or groups understand?

Openness and transparency

- How open can you be about this project? Could you publish your methodology, metadata, datasets, code or impact measurements?
- Can you ask peers for feedback on the project? How will you communicate it internally?
- Will you publish your actions and answers to this canvas openly?

Sharing data with others

- Are you going to be sharing data with other organisations? If so, who?
- Are you planning to publish any of the data? Under what conditions?



This is an adaptation of the Data Ethics Canvas produced by the ODI Learning team

Data Ethics Canvas

2020-06

DATA

IMPACT

ENGAGEMENT

PROCESS

Ongoing implementation

- Are you routinely building in thoughts, ideas and considerations of people affected by your project? How?
- What information or training might be needed to help people understand data issues?
- Are systems, processes and resources available for responding to data issues that arise in the long-term?

Reviews and iterations

- How will ongoing data ethics issues be measured, monitored, discussed and actioned?
- How often will your responses to this canvas be reviewed or updated? When?

Your actions

- What actions will you take before moving forward with this project? Which should take priority?
- Who will be responsible for these actions, and who must be involved?
- Will you openly publish your actions and answers to this canvas?



This is an adaptation of the Data Ethics Canvas produced by the ODI Learning team

Data Ethics Canvas

2020-06

DATA

IMPACT

ENGAGEMENT

PROCESS

Data sources

Name/describe your project's key data sources, whether you're collecting data yourself or accessing via third parties.
Is any personal data involved, or data that is otherwise sensitive?

Rights around data sources

Where did you get the data from? Is it produced by an organisation or collected directly from individuals?
Was the data collected for this project or for another purpose? Do you have permission to use this data, or another basis on which you're allowed to use it? What ongoing rights will the data source have?

Limitations in data sources

Are there limitations that could influence your project's outcomes?
Consider:

- > bias in data collection, inclusion/exclusion, analysis, algorithms
- > gaps or omissions in data
- > provenance and data quality
- > other issues affecting decisions, such as team composition

Ethical and legislative context

What existing ethical codes apply to your sector or project? What legislation, policies, or other regulation shape how you use data? What requirements do they introduce?
Consider the rule of law: human rights; data protection; IP and database rights; anti-discrimination laws; and data sharing, policies, regulation and ethics codes/frameworks specific to sectors (eg health, employment, taxation).

Ongoing implementation

Are you routinely building in thoughts, ideas and considerations of people affected by your project? How?
What information or training might be needed to help people understand data issues?
Are systems, processes and resources available for responding to data issues that arise in the long-term?

Your reason for using data

What is your primary purpose for collecting and using data in this project?
What are your main use cases? What is your business model?
Are you making things better for society? How and for whom?
Are you replacing another product or service as a result of this project?

Positive effects on people

Which individuals, groups, demographics or organisations will be positively affected by this project? How?
How are you measuring and communicating positive impact? How could you increase it?

Negative effects on people

Who could be negatively affected by this project?
Could the way that data is collected, used or shared cause harm or expose individuals to risk of being re-identified? Could it be used to target, profile or prejudice people, or unfairly restrict access (eg exclusive arrangements)?
How are limitations and risks communicated to people? Consider: people whom the data is about, people impacted by its use and organisations using the data.

Minimising negative impact

What steps can you take to minimise harm?
How could you reduce any limitations in your data sources? How are you keeping personal and other sensitive information secure?
How are you measuring, reporting and acting on potential negative impacts of your project?
What benefits will these actions bring to your project?

Reviews and iterations

How will ongoing data ethics issues be measured, monitored, discussed and actioned?
How often will your responses to this canvas be reviewed or updated? When?

Engaging with people

How can people engage with you about the project?
How can people correct information, appeal or request changes to the product/service? To what extent?
Are appeal mechanisms reasonable and well understood?

Communicating your purpose

Do people understand your purpose – especially people whom the data is about or who are impacted by its use?
How have you been communicating your purpose? Has this communication been clear?
How are you ensuring more vulnerable individuals or groups understand?

Openness and transparency

How open can you be about this project? Could you publish your methodology, metadata, datasets, code or impact measurements?
Can you ask peers for feedback on the project? How will you communicate it internally?
Will you publish your actions and answers to this canvas openly?

Sharing data with others

Are you going to be sharing data with other organisations? If so, who?
Are you planning to publish any of the data? Under what conditions?

Your actions

What actions will you take before moving forward with this project? Which should take priority?
Who will be responsible for these actions, and who must be involved?
Will you openly publish your actions and answers to this canvas?

Resources

theodi.org/dataethics

Data ethics canvas

Improve your practices

Data ethics whitepaper

Helping organisations navigate ethical concerns

MIT Moral Machine

<http://moralmachine.mit.edu/>

Kings Cross Facial Recognition

<https://www.ft.com/content/e9e57c2a-cff1-11e9-99a4-b5ded7a7fe3f>

Criminal Sentencing

<https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html>

Resources

theodi.org/dataethics

How children's lives have been affected by the pandemic

<https://theodi.org/article/new-report-data-about-childrens-lives-in-the-pandemic/>

Includes completed data ethics canvas for report

Please share your experience:

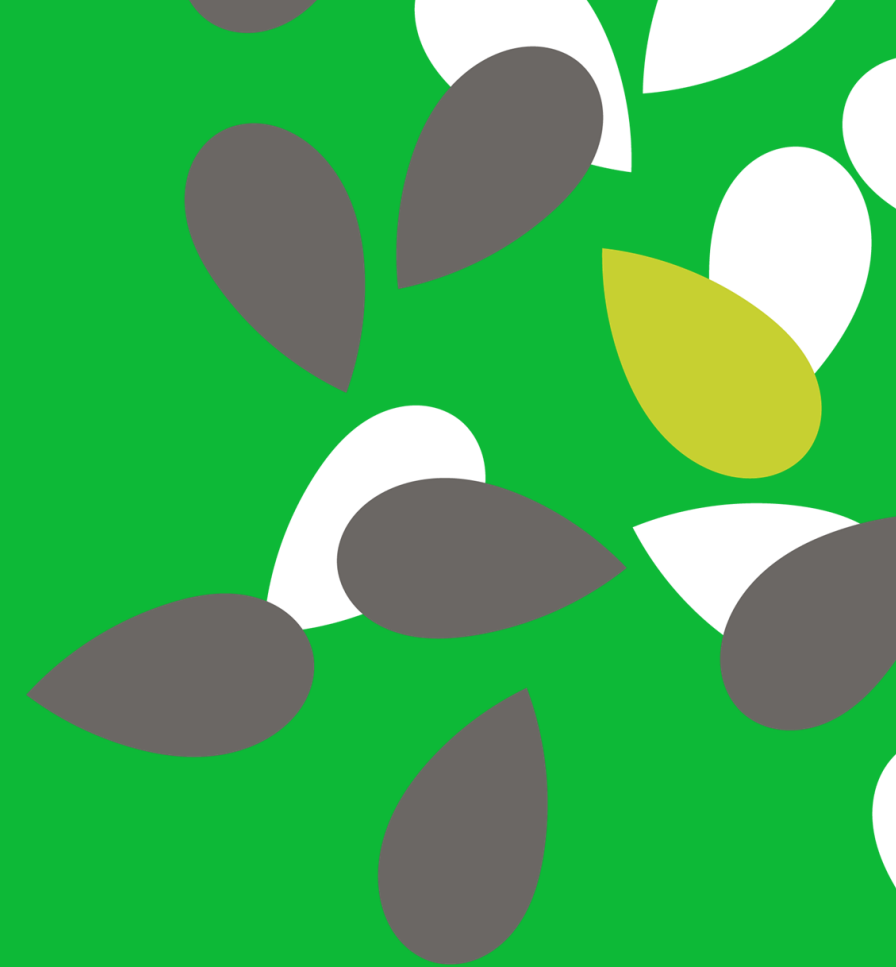
Feedback form

Get in touch

If you would like to talk to us about our training offering, or opportunities for collaboration, partnership, support or anything else, we'd love you to get in touch.

training@theodi.org

@ODIHQ



Next steps

Identify a data ethics challenge you would like to address.

Apply as much as you can of the data ethics canvas to the challenge.

Discuss and iterate this with colleagues.

Be careful to self identify the five stages of grief in yourselves!





Image citation: stevepb by Pixabay



nile by Pixabay



GoodGuyPaul by Pixabay

Car insurance

At the beginning of 2018, large global firms like Admiral and Marks & Spencers faced public backlash when the Sun newspaper found that insurance quotes for drivers with the traditional English name 'John' were far lower than quotes of the same for drivers named 'Mohammed'.

Adjust cover ✎

49 car quotes found, 4 telematics quotes included.

Sort: Annually ▾

Legal Assistance
+£30.99Breakdown Cover
+£43.99

Personal Accident



Windscreen



Courtesy Car

£538.26

Total Excess: £150

View details >

M&S BANK

Legal Assistance
+£26.29Breakdown Cover
+£31.54

Personal Accident



Windscreen



Courtesy Car

£571.66

Total Excess: £150

View details >

Legal Assistance
+£30.99Breakdown Cover
+£43.99

Personal Accident



Windscreen



Courtesy Car

£573.39

Total Excess: £150

View details >

Legal Assistance
+£30.99Breakdown Cover
+£43.99

Personal Accident



Windscreen



Courtesy Car

£577.87

Total Excess: £150

View details >

Adjust cover

37 car quotes found, 5 telematics quotes included.

Sort: Annually



Legal Assistance

+£30.99



Breakdown Cover

+£43.99



Personal Accident



Windscreen



Courtesy Car

£1,446.32

Total Excess: £150

View details >



Legal Assistance

+£30.99



Breakdown Cover

+£43.99



Personal Accident



Windscreen



Courtesy Car

£1,458.04

Total Excess: £150

View details >



Legal Assistance

+£30.99



Breakdown Cover

+£43.99



Personal Accident



Windscreen



Courtesy Car

£1,459.64

Total Excess: £150

View details >

M&S BANK



Legal Assistance



Breakdown Cover



Personal Accident



Windscreen



Courtesy Car

£1,476.70

Total Excess: £150

View details >

Assess the data ethics of a use case

What poses a higher risk to negative ethical impacts; the data **collection**, **sharing** or **use**?
And why?

+ What could have been done differently?

...

+ ...

Case 2: Criminal Sentencing

Link to the article ([here](#))
Please scroll to Frame 3 in Jamboard



COMPAS

Criminal Sentencing software.
Exploring the use of criminal records to aid judges make better decisions when predicting recidivism.