

# Ethical use of data

Version: 1.0



# Learning intentions

We will be looking at the **ethical use of data**, specifically

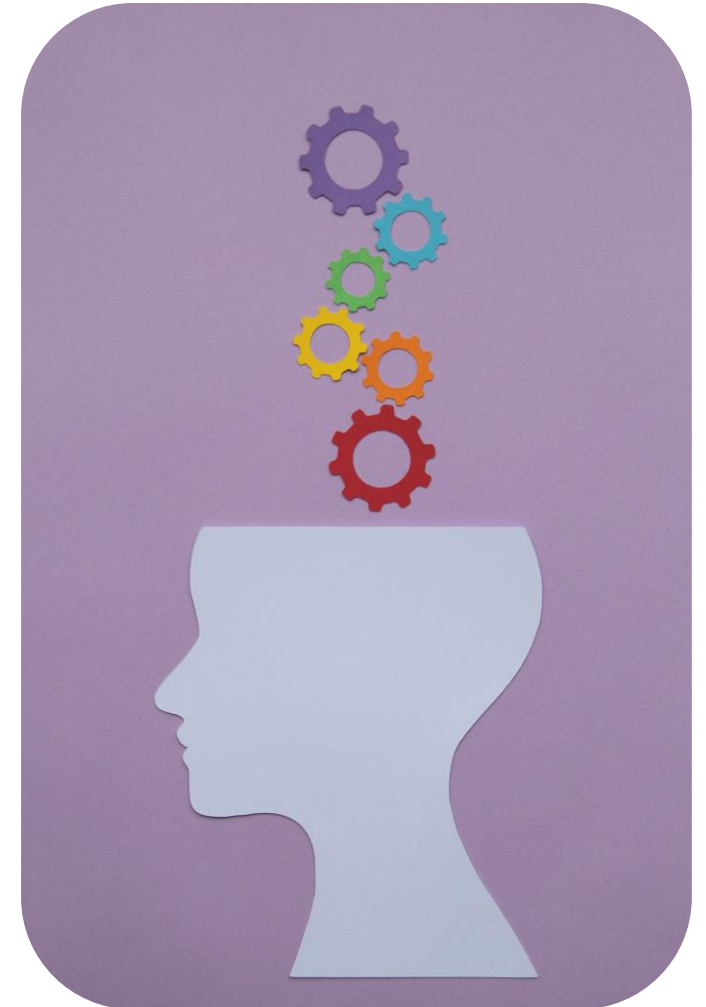
- What are ethical data risks
- How to identify ethical risks
- What are data ethics frameworks and how they can be used to prevent ethical risks

# Background

When processing data, people often jump straight to what they are going to do with it.

However, **reviewing the rights and wrongs of why you are processing the data is important.**

In this lesson, we are going to look the risks and problems that can arise from processing data.



Your turn...



What do you think of when someone says “**ethics**”?

Write down as many words as you can think of related to ethics  
(its ok if you can't think of any)



# Your turn...



What does **ethics** mean to you? Here are some example words, you might have more/different words

Trust  
Beliefs  
Rights  
Humane  
Fairness  
Honour  
Choices  
Principles  
Right vs. Wrong  
Responsibility  
Honesty  
Integrity  
Ideals  
Conscientious  
Morals  
Decency  
Standards  
Equality  
Behaviour

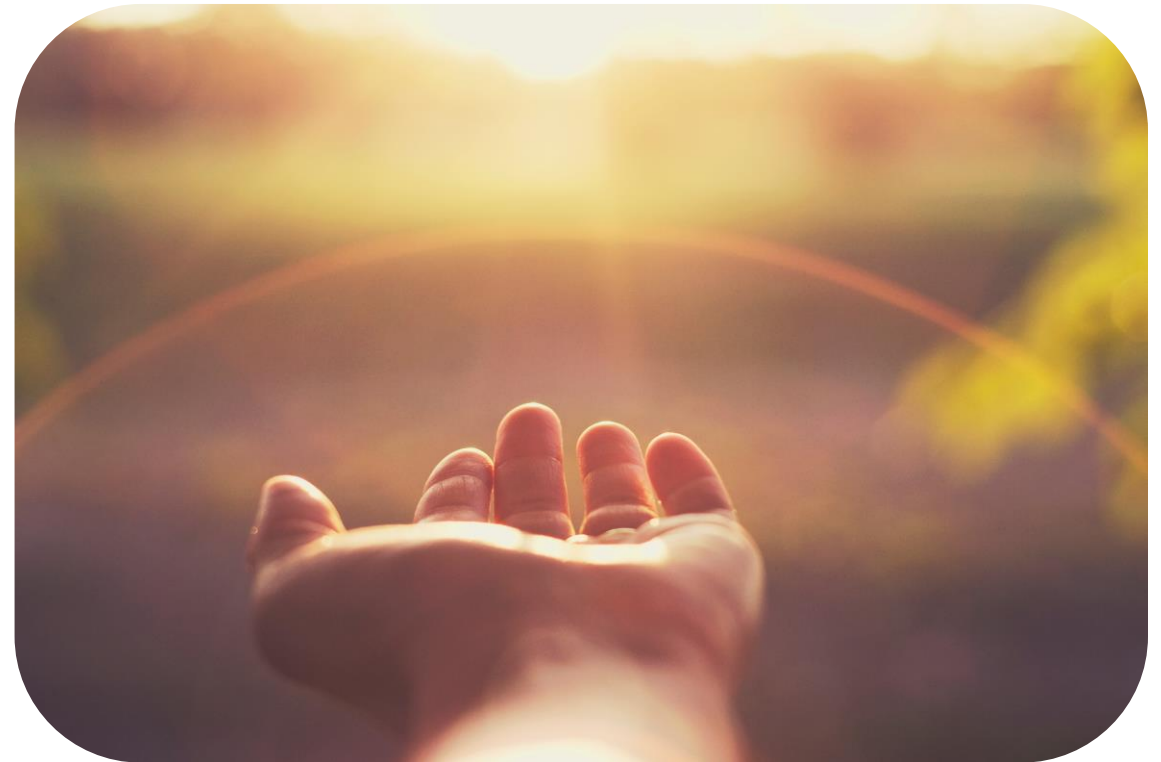
# What is “ethics”?

Ethics can be described in many different ways,

- **Set of rules** that helps us determine what is right and wrong.
- **Moral principles** that govern a person's behaviour
- A system of **accepted beliefs** that control behaviour.

In ethics, very rarely is there an absolute right or wrong answer.

This next video shows a problem that is often discussed related to ethics.







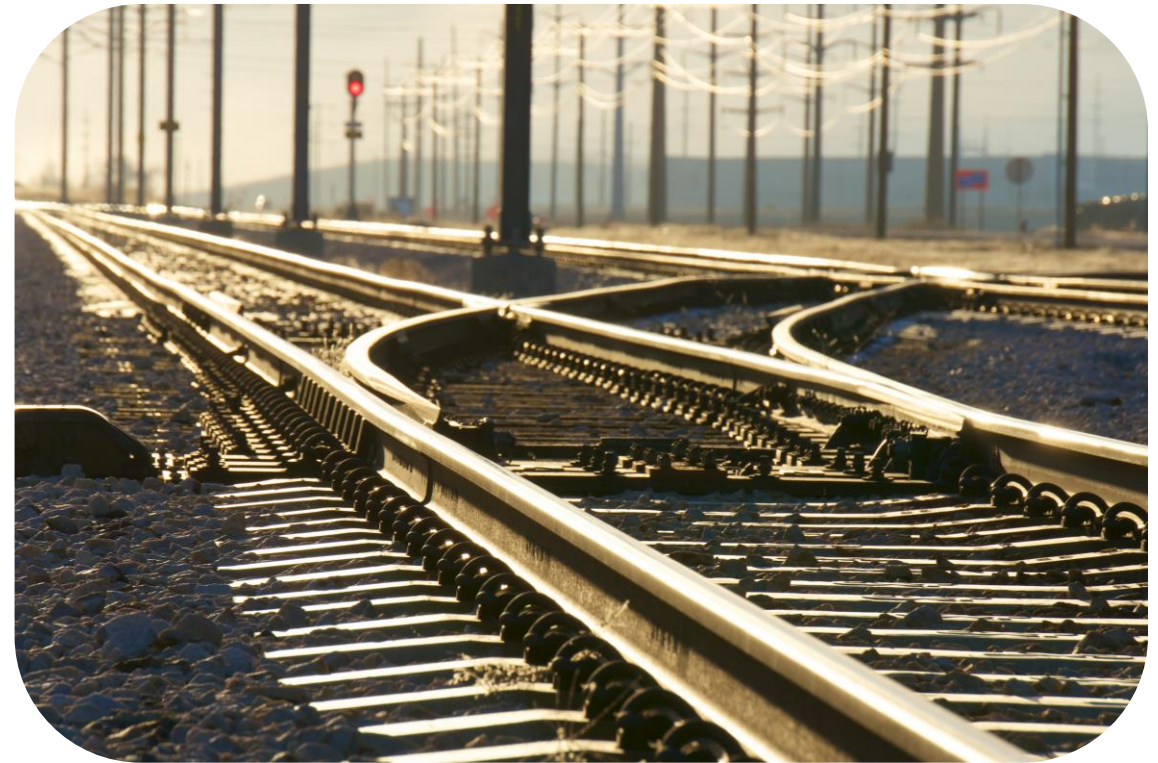
# The trolley problem

There is a **runaway trolley** barreling down the railway tracks towards **five people who are unable to move**.

You can pull a lever to change track. However, you notice that there is **one person on the side track**.

You have two (and only two) options:

1. **Do nothing**, in which case the trolley will kill the five people on the main track.
2. **Pull the lever**, diverting the trolley onto the side track where it will kill one person.





# Your turn...



Thinking about the trolley problem, do you think there **is a right answer** you could choose?

- Is **one option more ethical**? Does one more “do less harm”?
- If you **knew anything about the people on the track** (e.g. their age, if they were your friends), does that change the answer? **Is that ethical**?
- Do you think **everyone would make the same option choice** as you?



# Branches of ethics

Ethical choices rarely have an absolute right or wrong answer.

The results of **ethical choices impact every part of our lives**. Making sure the choices we are making are ethical can be complicated.

In this lesson, we are going to look the **ethical choices related to the use of data**.



Medical & research  
(bioethics)



Business ethics



Environmental  
ethics



Social ethics



Organisational  
ethics



**Data ethics**

# Definition



## **Data ethics**

Branch of ethics that looks  
at the moral questions  
related to data

# Show me...



Here are some examples of ways that data could be used, but are they ethical?



Police could use social media information to **predict which people are more likely to commit crimes.**



**Do we need elections,** or could we use polling data to elect politicians?



Do we need to sit exams or could we use **data to predict grades** instead?

# Data ethics risks and problems

The risks and problems relating to data ethics are wide and varied.

These risks and problems can be described by 8 areas.





# Data ethic risks and problems



## Truth

Technology could be used to either undermine the truth, or spread disinformation or propaganda. This may only be a risk if the technology falls into the wrong hands.



## Health

Technology could affect the mental, physical or social health of the users that engage with it by encouraging addiction or maximising user attention.



## Equality

Technology could be amplifying existing economic inequalities if those without access to it end up losing out. This could be through loss of employment, lack of knowledge or access to services.



## Fairness and Bias

Algorithms are often black boxes where it is not clear what is happening under the hood. The technology could be amplifying and reinforcing existing biases, having been developed on datasets containing historical bias.

# Data ethic risks and problems



## Human rights

Technology could infringe the human rights of citizens. This may be through increased surveillance or targeting of individuals leading to impacts on their future opportunities.



## Privacy

Personal data collected could be being disclosed with the individual's knowledge or control. This may lead to companies profiting from individual's data without their knowledge.



## Trust and Transparency

User may not be fully aware of how their data is being used within the technology. This could lead to a backlash from users as their trust is eroded



## Criminality

Technology could be used for illegal activities, even if it were not designed for this.



# Show me...



In China they use **The Social Credit System**. This brings together financial and social assessments of businesses, government institutions and individuals.

The aim of the system was to **build trust** and regulate areas such as food safety and financial fraud. It also helps regulate social behaviour (e.g. paying bills on time) and **promotes traditional moral values**.

However, it adds in information from the [SkyNet](https://en.wikipedia.org/wiki/SkyNet) system which is a **network of c.200 million CCTV cameras** which uses facial recognition to track everything individuals are doing. Anything that is seen as suspicious (e.g. leaving the house by the back door) is highlighted in the system.



# Your turn...



Which of the **data ethical risks/problems** below could arise when thinking about the **mass surveillance in China**?

- Truth
- Health
- Equality
- Fairness and Bias
- Human Rights
- Privacy
- Trust and Transparency
- Criminality



# Your turn...



Which of the data ethical risks/problems below could arise when thinking about the mass surveillance in China?

- **Truth** – the technology could be used to undermine truth or spread propaganda
- Health
- **Equality** – the technology could amplify existing economic inequalities
- **Fairness and Bias** – could amplify and reinforce existing bias
- **Human Rights** – could infringe the human rights of citizens
- **Privacy** – the personal data could be shared without the individual's control
- **Trust and Transparency** – the user may not be fully aware of how there is being used
- Criminality

# Show me...



In 2018, Amazon had been developing a tool that could **review CVs submitted** for jobs and then **decide who they should hire**.

The tool used applications submitted over a 10-year period, most of the applications were from men.

As the tool was built on a dataset dominated by male applicants, the **system taught itself that male candidates were preferable**. It penalised CVs that included the word “women” as in “women’s football team captain”.

Amazon stopped the project and said it “was never used by Amazon’s recruiters to evaluate candidates”.





# Your turn...



## What do you think of the Amazon recruitment tool?

- What types of ethical risks/problems did the tool have?
- Could they have overcome the risks/problems?
- Would you like this tool to be used for your job application?
- Do you think the tool is more or less fair than a human interviewer?



*Remember there are rarely absolute right or wrong answers in ethics.*

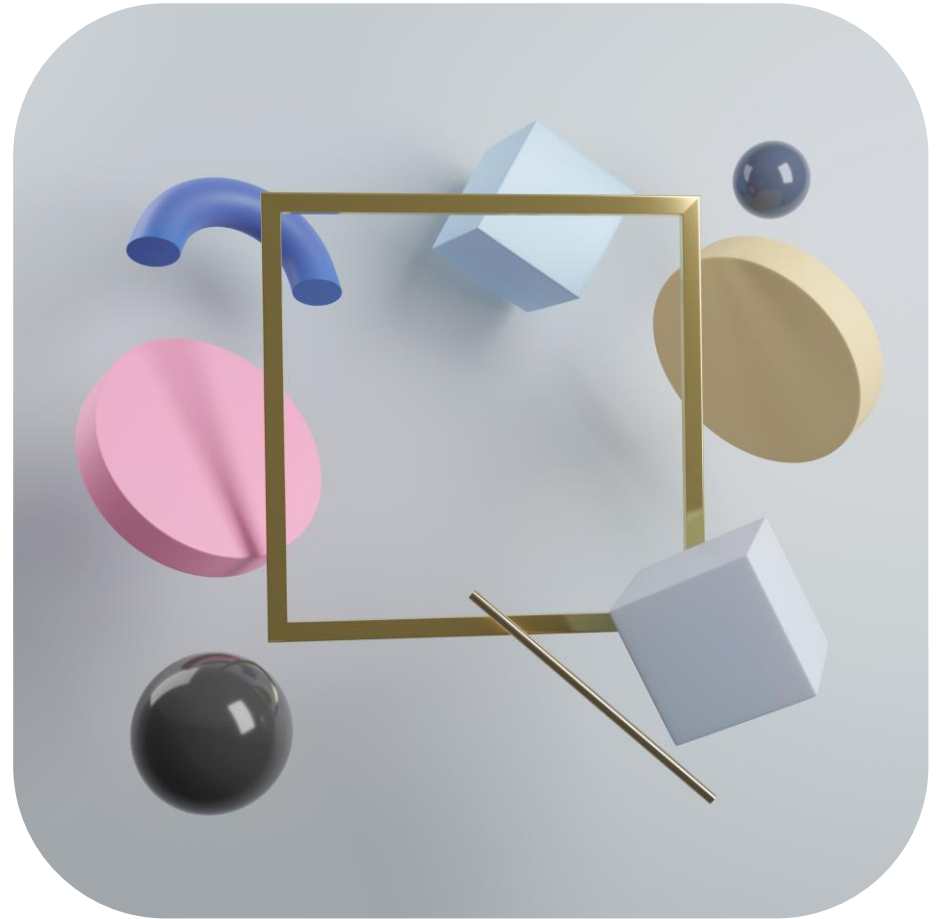
## Next steps

Complete **questions 1 to 5**  
in **section 1** of the  
'Ethical use of data' workbook.

# Frameworks for data ethics

As we have seen the issues related to data ethics can be wide and varied.

To help individuals and companies to use data in an ethical way, frameworks have been developed that you can follow.





# Definition



## **Data ethics frameworks**

Ensures that choices related to the use of data within an organisation reflect and uphold its ethical principles

# Show me...



Frameworks often contain questions you can ask yourself as you go through processing of data.



## 5.6 Public scrutiny (accountability)

- Do members of the public or end users of the project have the ability to raise concerns and place complaints about the project? If yes, how? If not, why?
- What channels have you established for public engagement and scrutiny throughout the duration of the project?

## 5.7 Share your learnings (transparency)

How have you documented and shared the progress and case studies from your project with peers and stakeholders?

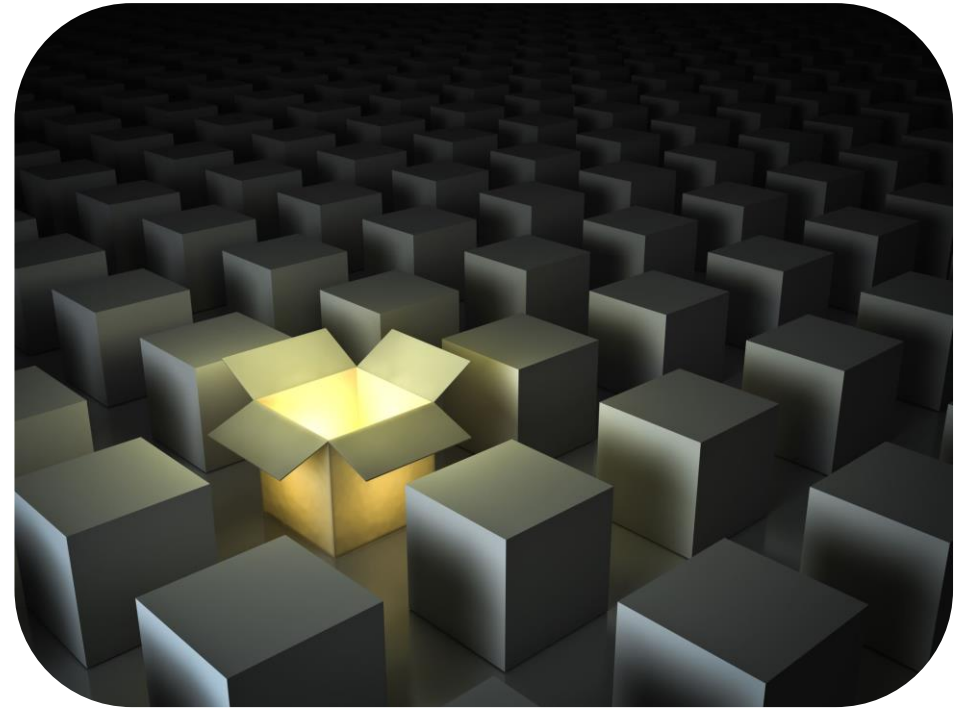
## Next steps

If you have scored a 3 or less in any of the principles, this could indicate the need for additional checks and potential changes to make your project more ethical. Please explain the reason for the score and consult the outcome with your team leader, organisational ethics board or data ethics lead to advise on the specific next steps to improve the ethical standards of your project.

# Frameworks and its impacts

Frameworks can have a positive impact on the organisations/companies that use them. These can include,

1. Ensuring **different types of ethical risks are reviewed**
2. Ensuring **fairness**
3. Ethical **standards are consistent** across the organisation
4. Helps with **openness and transparency**
5. Ensures they **comply with the law**
6. Helps the **prevention of harm**



# Framework questions

Frameworks can be helpful when thinking about data ethics, however they should not be used as a “tick box” or a “one-off” exercise.

Frameworks can help you to think about,

- Is there...?
- What if....?
- Should I...?





# Example

A cinema want to understand who their “loyal customers” are. They have access to lots of data including banking information and the films customers have watched.

Here are some examples of ethical data questions,

- **Is there** consent from the customers to process the data like this?
- **What if** processing the data impacts of the trust of the customers?
- **Should** the films customers have watched be analysed?



# UK government data ethics frameworks

UK Government has a Data Ethics Framework for using data within the government and the wider public sector.

Each section contains detailed questions which are scored between 0 to 5. Any question scoring 3 or less, needs additional checks and possible changes to make the project more ethical.

## Areas of the framework

1. Define and **understand public benefit** and user need
2. Involve **diverse expertise**
3. **Comply with the law**
4. Review the **quality and limitations of the data**
5. Evaluate and consider **wider policy implications**

# ‘Responsible Data for Children’ principles

The organisation ‘Responsible Data for Children’ has 7 principles it follows to make sure it is using data ethically.

1. Participatory
2. Professional accountable
3. People-centric
4. Prevention of harms across the data life cycle
5. Proportional
6. Protective of children’s rights
7. Purpose driven





# Keeping frameworks relevant

By using a data ethics framework, it allows everyone in an organisation to see how they are expected to use the data.

**Frameworks should comply with regulations** (such as GDPR) and be **regularly reviewed** to make sure they are keeping up with technological and society changes.



## Next steps

Complete **questions 1 to 5**  
in **section 2** of the  
'Ethical use of data' workbook.

# Learning checklist

I can *explain* what is meant by data ethics

I can *explain* what are data ethical risks

I can *explain* how to identify data ethical risks

I can *explain* what is a data ethics framework and how it protects data

# How you can use this lesson



You are free to:

- **Share** – copy and redistribute the material in any medium or format
- **Adapt** – remix, transform and build upon the material

Under the following terms:

- **Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **NonCommercial** — You may not use the material for [commercial purposes](#).
- **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.

© 2022. This work is licensed under a [CC BY-NC-SA 4.0 license](#).

Created by effini in partnership with Data Education in Schools and Skills Development Scotland.



# Alternative format

If you require this document in an alternative format, such as large print or a coloured background, please contact

**hello@effini.com**

or

**4th Floor, The Bayes Centre  
47 Potterrow  
Edinburgh  
EH8 9BT**



**Skills  
Development  
Scotland**