Queen Mary
University of London
Science and Engineering

ECS7025P Ethics, Regulations and Laws in Advanced Data Processing and Decision Making

Week 3

Specific Action points – Part A

Dr Mahesha Samaratunga

Learning Objectives:

- Introduction to the first specific action in UK Gov's Data Ethics Framework
- Exploring and understanding the public benefit and user needs
- Apply the sub-steps in the specific action to real life examples.

What are the potential benefits and need for user

- The first step is to **Define and understand the benefits to the society and the needs of the user**
- All successful projects are dependent on a clear problem definition, and the needs can help you define the problem that you are trying to solve.
- You also need to see if the data science project you are conducting will lead to benefits for the society.

1. Understand the wider public benefit

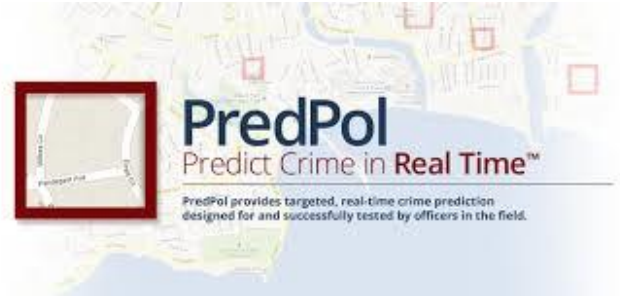
- What are direct benefits for individuals in this project? (for example saving time when applying for a government service)
- How does the project deliver positive social outcomes for the wider public?
- How can you measure and communicate the benefits of this project to the public?
- What are the groups that would be disadvantaged by the project or that would not benefit from the project? What can you do about this?

2. Understand unintended consequences of the project

- What would be the harm in not using data? What social outcomes might not be met?
- What are the potential risks or negative consequences of the project versus the risk in not proceeding with the project?
- Could the misuse of the data or algorithm or poor design of the project contribute to reinforcing social and ethical problems and inequalities?
- What kind of mechanisms can you put in place to prevent this from happening?
- What specific groups benefit from the project? What groups can be denied opportunities or face negative consequences because of the project?

The criminal minority

- Already in use in several US states, PredPol is an algorithm designed to predict when and where crimes will take place, with the aim of helping to reduce human bias in policing.
- But in 2016, the Human Rights Data Analysis Group found that the software could lead police to unfairly target certain neighbourhoods. When researchers applied a simulation of PredPol's algorithm to drug offences in Oakland, California, it repeatedly sent officers to neighbourhoods with a high proportion of people from racial minorities, regardless of the true crime rate in those areas.




THOMSON REUTERS
FOUNDATION
A
ON THE EDGE
Policing
Algorithms

- In response, PredPol's CEO pointed out that drug-crime data does not meet the company's objectivity threshold, and so in the real world the software is not used to predict drug crime in order to avoid bias.
- Even so, last year Suresh Venkatasubramanian of the University of Utah and his colleagues demonstrated that because the software learns from reports recorded by the police rather than actual crime rates, PredPol creates a "feedback loop" that can exacerbate racial biases.

3. Human rights considerations

- How does the design and implementation of the project or algorithm respect human rights and democratic values?
- How does the project or algorithm work towards advancing **human capabilities**, advancing **inclusion of underrepresented populations**, reducing economic, social, gender, racial, and other **inequalities**?
- What are the **Environmental limplications** of the project? How could they be mitigated?

4. Justify the benefit for the taxpayers and appropriate use of public resources in your project

- How can you demonstrate the value for money of your project?
- Is there effective governance and decision-making oversight to ensure success of the project?
- Do you have evidence to demonstrate all of the above?

5. Make your user need and public benefit transparent

- Where can you publish information on how the project delivers positive social outcomes for the public?
  - Considering the private sector
- How have you shared your understanding of the user need with the user?

6. Understand the user need

‘User needs’ are the needs that a user has of a service, and which that service must satisfy for the user to get the right outcome for them.

Example:

- running and improving services
- building new services
- trialling new processes for internal operations
- testing existing and new policies

7. Ensure there is a clear articulation of the problem before you start the project.

Do you understand the problem well?

Do you understand the solution well?

How will the solution benefit the user?

What are the implications of using the solution to users?

8. Check if everyone in your team understands the user need and how using data can help

- Does everyone in your team understand the user need?
- Often projects involving data analysis are requested by non-practitioners - people with an ill-defined problem they would like to understand better.
- Reframing their request as a user need will help you understand what they're asking for and why, or expose what you don't know yet.

Try these things to understand better..

- Keep researching throughout each development phase to make sure the project to meet user needs.
- Research any past similar projects
- Consult more widely or get independent advice
- Speak to operational staff to understand how data is collected and what impact any new tool might have

Queen Mary
University of London
Science and Engineering

ECS7025P Ethics, Regulations and Laws in Advanced Data Processing and Decision Making

Specific Action 2: Involve Diverse Expertise

Dr Mahesha Samaratunga

Learning Objectives:

- Introduction to the 2nd specific action in the UK Gov's Data Ethics Framework

Involve diverse expertise

- Working in diverse, multidisciplinary teams with wide ranging skill sets contributes to the success of any data or tech project.

- If you do not have the sufficient skills or experience, you should involve others from your team or wider network with the right expertise.

Get the right expertise

- Beyond data scientists, who are the relevant **policy experts and practitioners** in your team?
- What **other disciplines and subject matter experts** need to be involved?
  - What steps have been taken to involve them?
- What should their **roles** be? Have you defined who does what in the project?
- Ask your team and experts if you **have the right data** for your research questions.

Ensure diversity within your team (fairness)

- How have you ensured **diversity** in your team?
  - Having a diverse team helps prevent biases and encourages more creativity and diversity of thought.
  - Diversity in: gender, age, ethnicity, religious belief, interest, etc.

- **Avoid** forming **homogenous** teams

Involve external stakeholders

- How have you engaged **external domain experts** in your project
  - E.g. academics, ethicists, researchers
  - Research communities
  - Government networks
  - Private sector organizations, funders, consultants, and technology-driven companies
- Relevant **civil society organisations**?
- Consider consulting the **target audience or the users** of your project?

Potential actions to build capacity…

- Appointing **chief data/AI officers** with ethics as part of their responsibilities
- Assembling organizationally high-level **ethics advisory groups**
- Incorporating privacy and **ethics-oriented risk and liability assessments** into decision-making or governance structures
- Providing **trainings and guidelines on responsible data practices** for employees

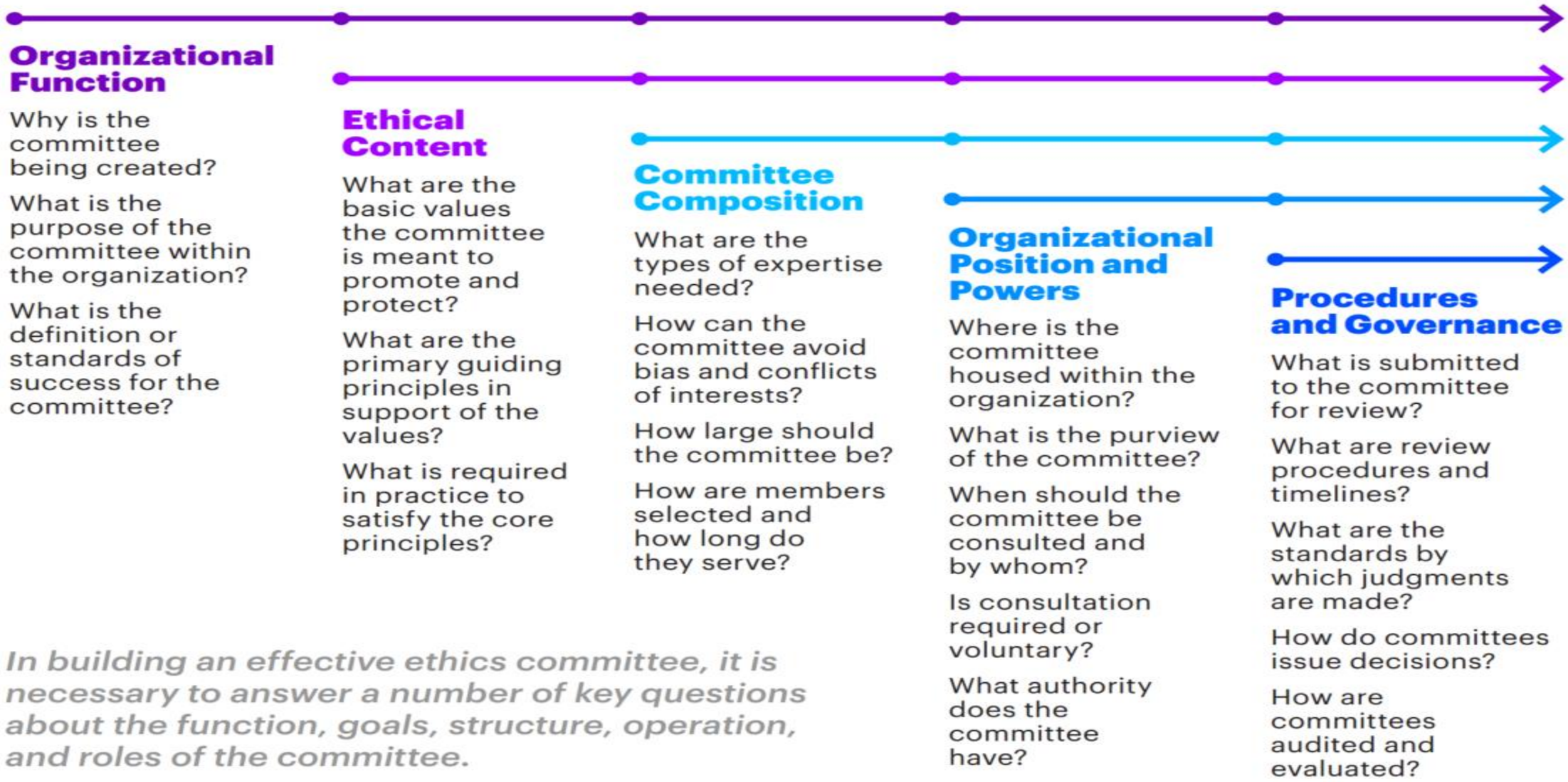
Potential actions to build capacity…

- **Developing tools, organizational practices/structures, or incentives** to encourage employees to identify potentially problematic data practices or uses

- Including members responsible for representing **legal, ethical, and social** perspectives on technology research and project teams

- Creating **ethics committees** that can provide guidance not only on data policy, but also on concrete decisions regarding collection, sharing, and use of data and AI

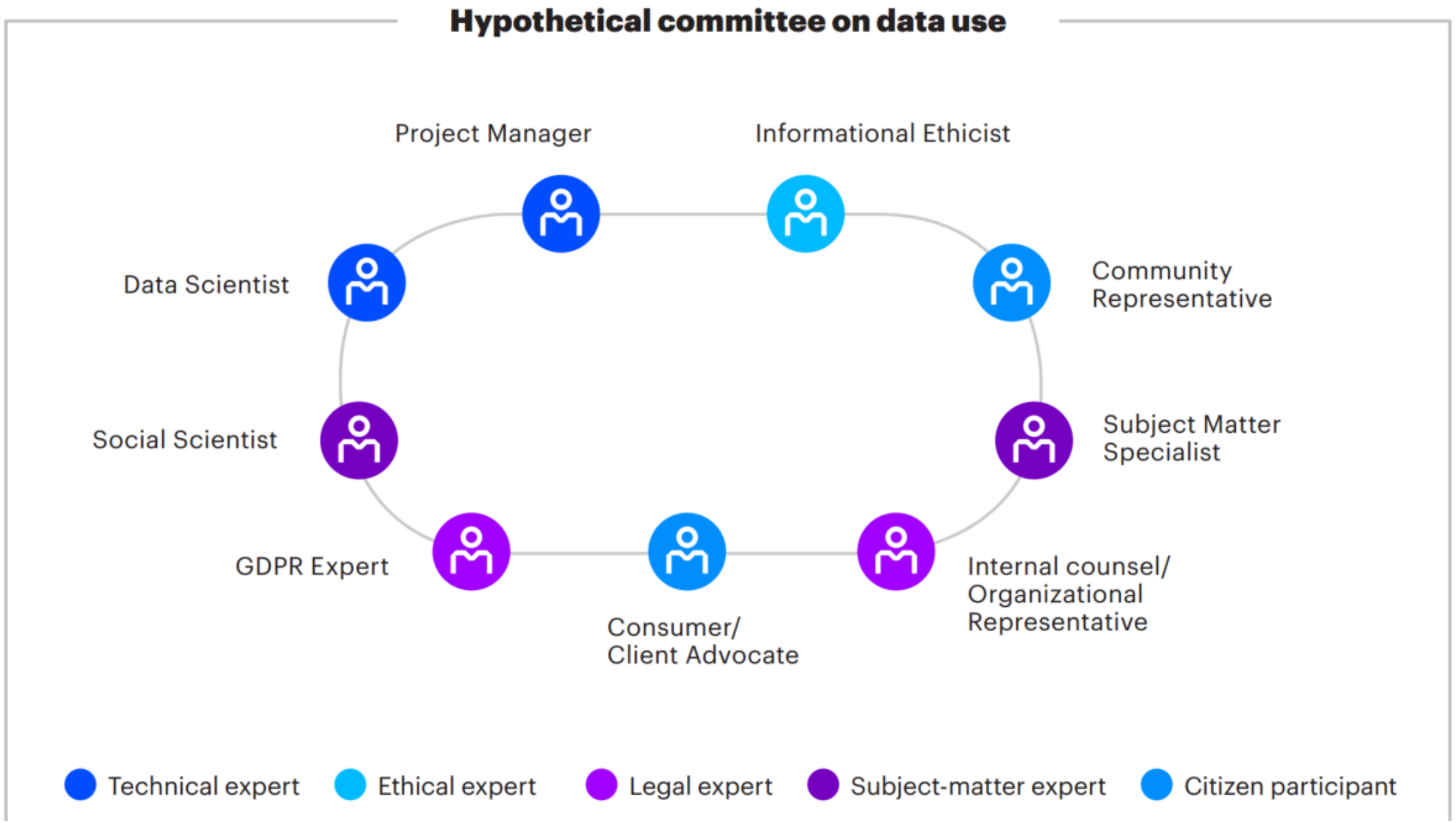
A roadmap for building an ethics committee



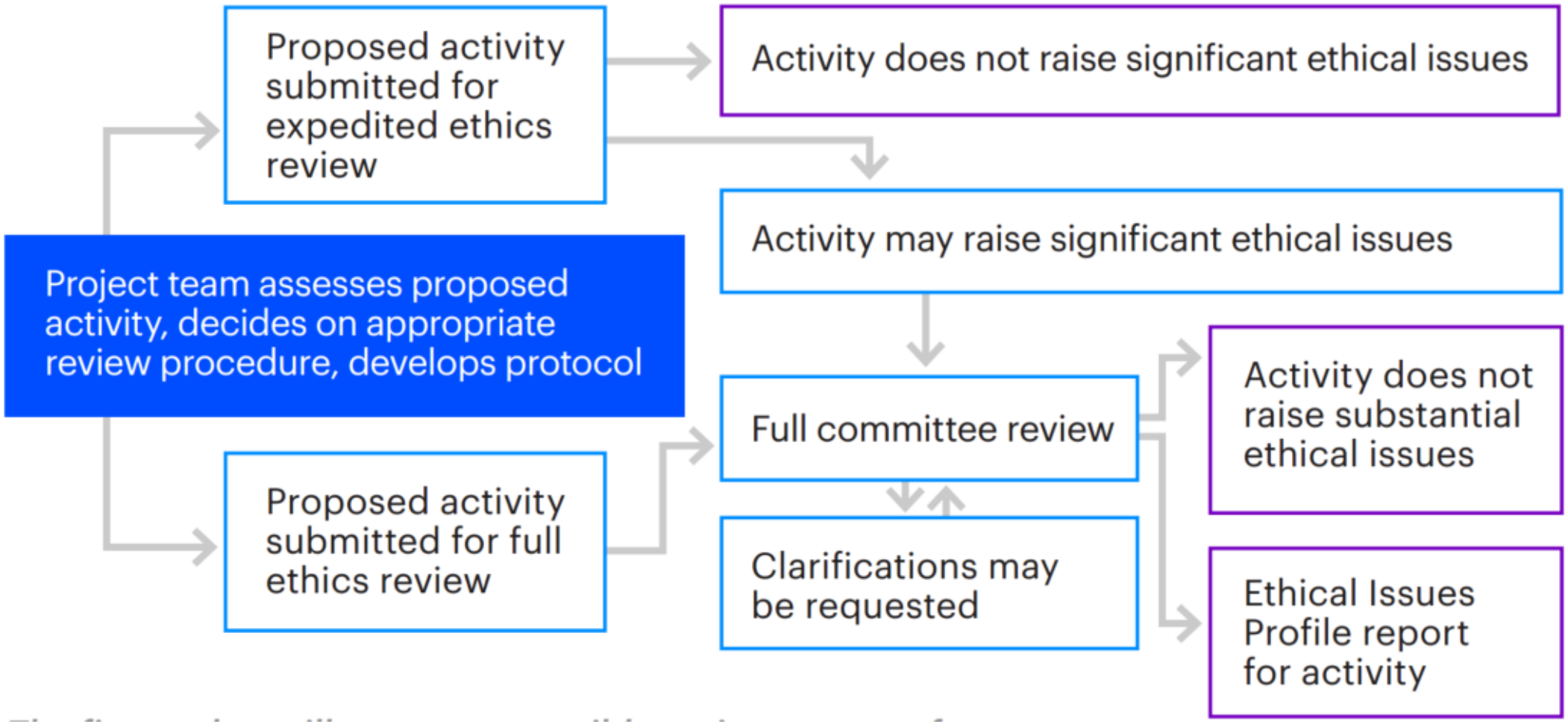
In building an effective ethics committee, it is necessary to answer a number of key questions about the function, goals, structure, operation, and roles of the committee.

Committee Composition

- Technical Experts
- Ethical Experts
- Legal Experts
- Subject Matter Experts
- Citizen Partcipants


Hypothetical committee on data use
Project Manager
Informational Ethicist
Data Scientist
Community Representative
Social Scientist
Subject Matter Specialist
GDPR Expert
Internal counsel/ Organizational Representative
Consumer/ Client Advocate
Technical expert
Ethical expert
Legal expert
Subject-matter expert
Citizen participant

Potential ethics committee review process



Queen Mary
University of London
Science and Engineering

ECS7025P Ethics, Regulations and Laws in Advanced Data Processing and Decision Making

Specific Action 3: Comply with the Law

Dr Mahesha Samaratunga

Learning Objectives:

- Introduction to the 3rd specific action point in UK Gov's Data Ethics Framework—Comply with the law

Get legal advice

- Have you spoken to a legal adviser within your organisation?
- Have you spoken to your information assurance team?
- Have you consulted your organisation's Data Protection Officer when doing a DPIA?
- What legal advice have you received?

What data are we allowed to use?

As a data scientist you will be making use of a wide range of datasets to conduct your analysis. This will also involve quantitative secondary research sources that includes data such as census data, birth/death rates, unemployment rates. This type of data is normally generated by governments, organisations and charities. Which leads us to the question: *Are we allowed to make use of this data?* The answer is ‘yes’, however you need to be aware of legislations related to the usage of data. According to gov.uk, This includes how you:

- produce statistics,
- protect privacy by design,
- minimise the data needed to achieve your need,
- keep personal and non-personal data secure.

Personal data Protection

If you are using personal data, you must comply with the principles of the

- EU General Data Protection Regulation (GDPR) and
- Data Protection Act 2018 (DPA 2018) which implements aspects of the GDPR and transposes the Law Enforcement Directive into UK law.

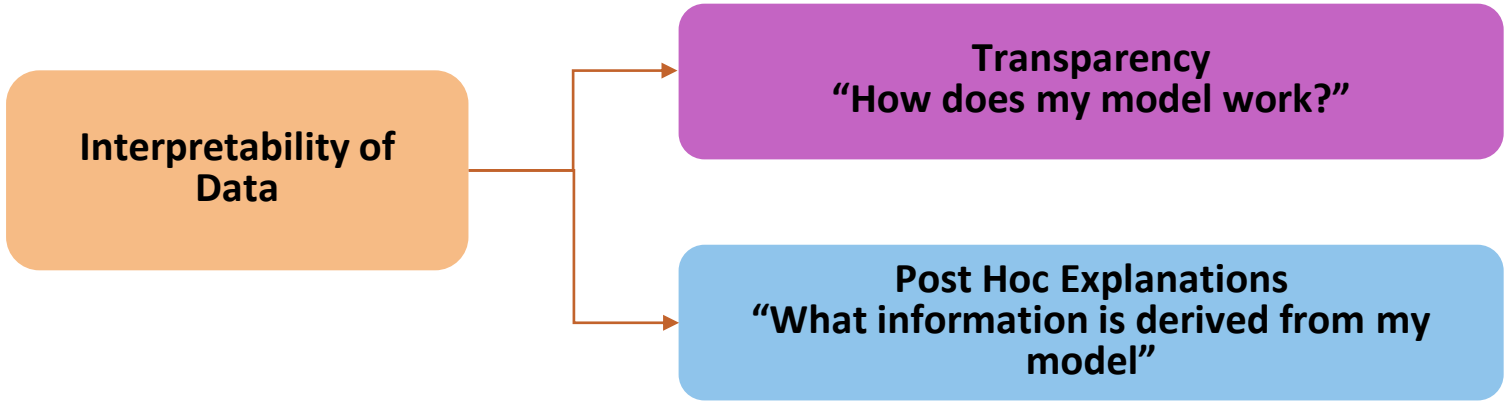
Personal data is defined in Section 3(2) DPA 2018 (a wider explanation is detailed in Article 4 of the GDPR).

- "Personal data" means any information relating to an identified or identifiable living individual
- "'personal data' means any information relating to an identified or identifiable natural person ('data subject')……; " --- Article 4 of GDPR

A Summary of EU GDPR

Data Protection Act 2018

Data scientists also need to take into consideration the interpretability of data, as this is also a GDPR requirement. There are two aspects to the interpretability of data, which are transparency and post hoc explanations. Transparency is based on how your model works, while post hoc explanations are based on the information derived from your model. From a GDPR perspective, this is important as a user has the legal right to find out how an algorithmic decision was made about them.



Further Reading

After class, you should read the relevant legislation documents to familiarise yourself with the legal implications:

Producing statistics - Code of Practice for Statistics

Minimising the data needed to achieve your need - Article 5(1)(c) of GDPR

Information governance - keeping personal and non-personal data secure (e.g. collection, storage, sharing and deletion) - General Data Protection Regulation (GDPR) and the Data Protection Act 2018 (DPA 2018)

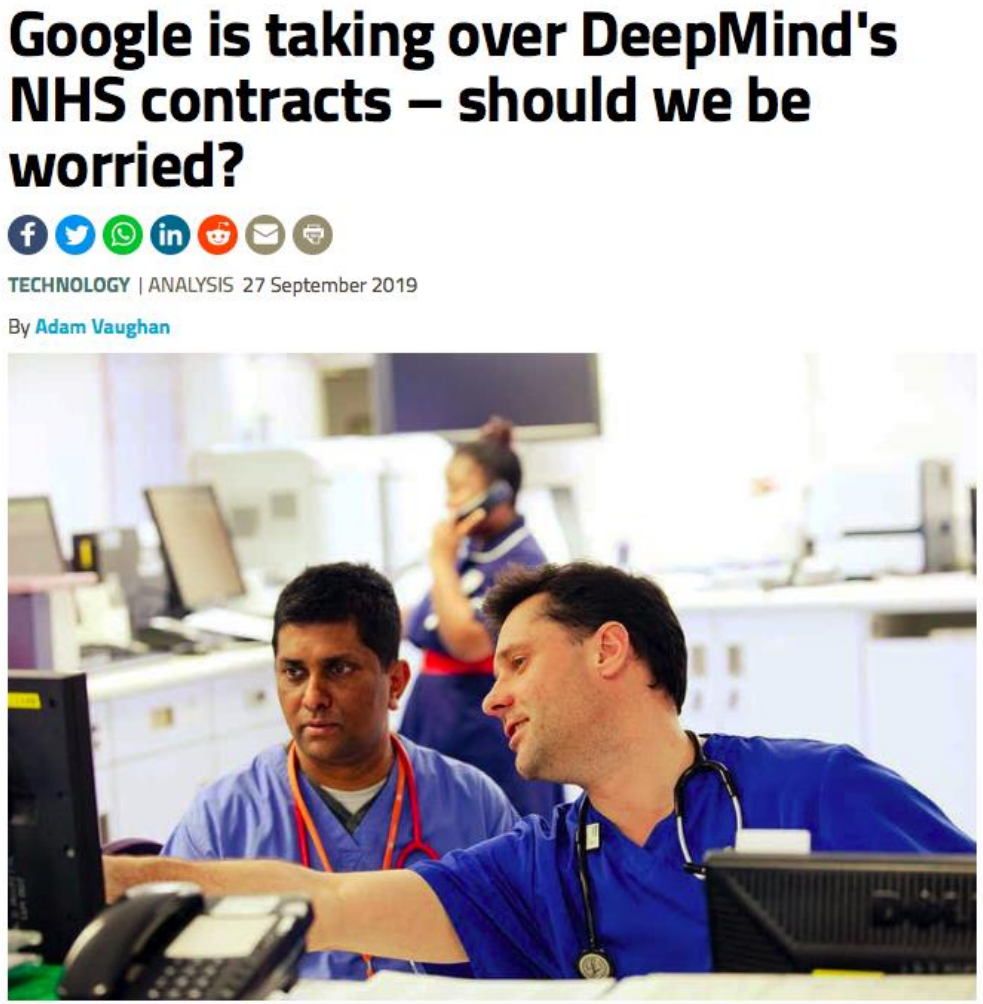
- GDPR Enforcement Tracker

GDPR case study

Read the following article and discuss the following questions:

https://www.newscientist.com/article/2217939-google-is-taking-over-deepminds-nhs-contracts-should-we-be-worried/

- Is this a privacy concern? If yes, then why?
- Are the patients obliged to share their data?



Google is taking over DeepMind's NHS contracts – should we be worried?

TECHNOLOGY | ANALYSIS 27 September 2019

By Adam Vaughan

Data protection by design and DPIA

GDPR requires that anyone handling personal data protects the rights of individuals by:

- Using personal data for a specific task
- Putting in place technical and organisational measures to implement data protection principles effectively
- Integrating necessary safeguards into the processing of personal data

It is a legal obligation under Article 35 of the GDPR to complete a **DPIA (Data Protection Impact Assessment)** when there's likely to be high risk to people's rights, particularly when using new technologies. However it is often good practice to do a DPIA for any use of personal data.

The DPIA template

Things to think about:

- Always seek the advice of your organisation's Data Protection Officer when doing a DPIA
- Privacy should be considered throughout the project – although you may not be using personal data at the outset of your work, the project type and privacy considerations may change as work develops
- Consider how often you will repeat the DPIA when using personal data and may need to change this if the project changes significantly
- When joining a new project, seek out and review the existing DPIA to familiarise yourself with any risks to rights and freedoms identified and the relevant mitigation strategies proposed
- If you discover a DPIA has *not* been completed for a project for which it is relevant, this should be flagged as soon as feasible
- Refer to the ICO's guidance on DPIAs

Legal Liability

Decision making models are dependent on data that is generated given a particular scenario. One such example is the series of decisions that have to be made given the data captured by the multiple sensors in Autonomous Vehicles (AVs). The question that we need to think about are:

“Who makes these decisions?”

“Are there any legal liabilities for these decisions?”

Scenario 1

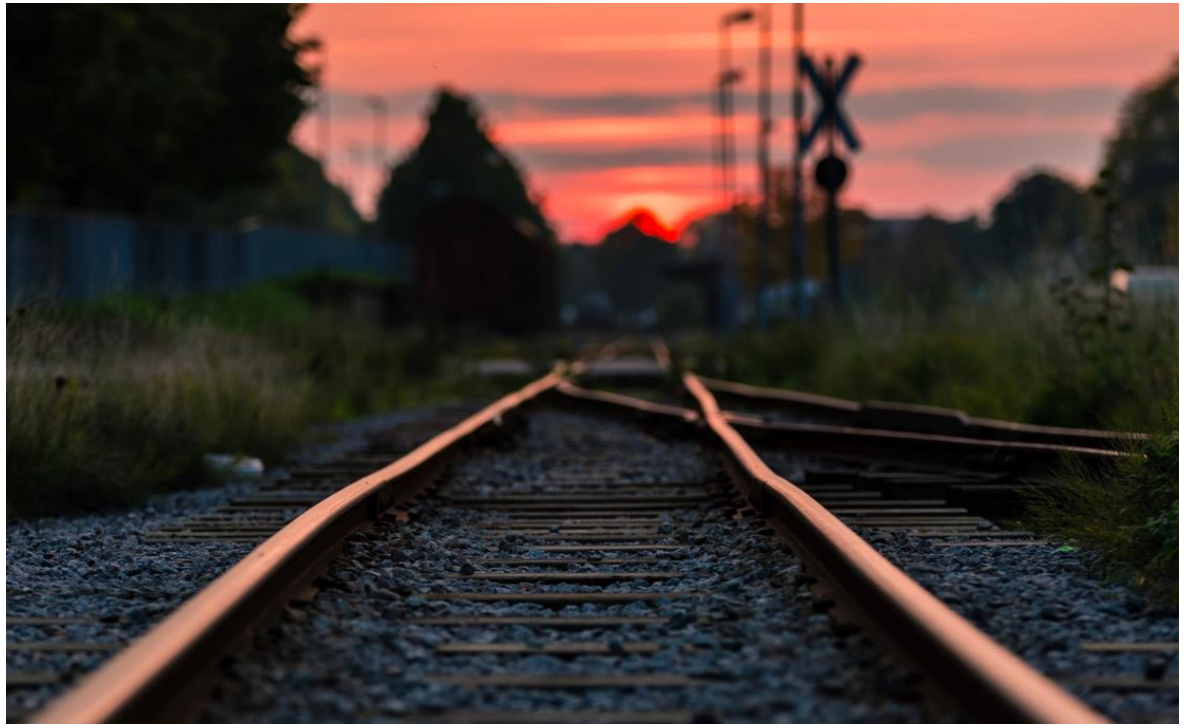
Imagine a runaway train is madly hurtling towards 5 people on a railway track – you look around but there's no way of warning them.

You notice that you're standing next to a lever that operates some points – you can divert the train onto another track and save the people – hurrah!

But there is a problem – there is also a person on the other track. If you hit the lever and divert the train then they die.

If you do nothing then 5 people die, but if you pull the lever then 5 people are spared but 1 person dies.

What would you do?



Scenario 2

Now let's change some of the conditions and look at this again. Imagine the train heading for those 5 people again. This time though, there are no points – but there is a very very large person on a bridge above the tracks.

You realise that the only way of saving the 5 people is to push the large person onto the tracks, just as the train approaches. They would die, of course, but the 5 would be spared. I mean, you have to suspend some disbelief here – the person would have to be huge to stop a train!

Similar ethical dilemmas happens to designing autonomous vehicles algorithm.

Will the legal liabilities be the same for a pre-programmed AV and a human-driven car?

Accountability & Transparency

Article 5(2) of the GDPR says:

> "The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 ('accountability')

Art. 30 GDPR

Records of processing activities

1. [1] Each controller and, where applicable, the controller's representative, shall maintain a record of processing activities under its responsibility. [2] That record shall contain all of the following information:

Publish your DPIA and other related documents

Equality and discrimination

Analysis or automated decision making must not result in outcomes that lead to discrimination as defined in the Equality Act 2010.

Information governance

- Organisations have a responsibility to keep both personal data and non-personal data secure.

- In many organisations information risk is (should be) overseen by a Senior Information Risk Owner (SIRO). Usually, your organisation will have a risk appetite statement that sets out how information risk is managed.

- Consult with your information assurance team when you need to delete data.

Sharing and re-use of data

- When accessing or sharing personal data, you must follow the Information Commissioner's Code of Practice for Data Sharing which should be read alongside the ICO's guide to GDPR. This code of practice has been updated in December 2020.
- When accessing and sharing data under powers in Part 5 of the Digital Economy Act 2017, you must follow the relevant codes of practice.
- When re-using published and unpublished information relating to public tasks, you must follow the Re-use of Public Sector Information Regulations 2015.

Copyright and intellectual property

Copyright and (Intellectual Property (IP) are often governed by combinations of statutes.

- When using data, respect copyright laws and database rights, covered in part by the Copyright and Rights in Databases Regulations 1997.

- When procuring software, consider potential intellectual property constraints covered in the Intellectual Property Act 2014.

Freedom of information

- The use of data may be subject to the Freedom of Information Act 2000.

- Also consider the wider publishing of datasets released following a Freedom of Information request, in accordance with the Protection of Freedoms Act 2012.

Sector specific legislation

Specific sectors like finance and health have further data use legislation and frameworks, including those relating to the use of non-personal data.

- Health research has its own UK Policy Framework for Health and Social Care Research drafted by the NHS Health Research Authority (HRA).
- The NHS HRA also provides specific guidance for health researchers on the new data protection principles being introduced by the GDPR.

Other regulations, laws and code of practices in UK

Other important pieces of central government guidance that are helpful for using data and designing projects in the public sector include:

- The Civil Service code
- HM Treasury Aqua Book: guidance on producing quality analysis for government
- HM Treasury Magenta Book: guidance for evaluation

Queen Mary

University of London