

A graphic element on the left side of the slide features a yellow circle containing a lit lightbulb. Three black lines radiate outwards from the top of the bulb, suggesting light or an idea. The lightbulb is connected to a black cord that extends downwards.

# Common Ethical Challenges

for Data Practitioners and Users

*\* based on Introduction to Data Ethics module (Part 2)*

# 1. Ethical Challenges in Appropriate Data Collection and Use

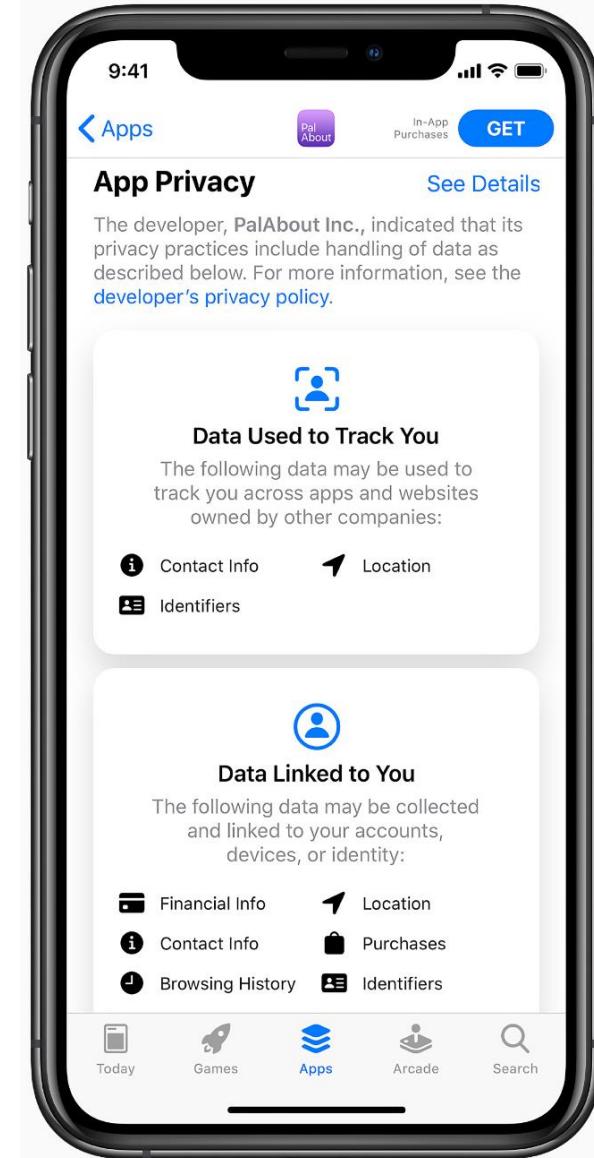
How can we properly acknowledge and respect the **purpose** for, and **context** within which, certain data was shared with us or generated for us?

- In a **medical** context:
  - A patient may share their medical records with their doctor.
  - A doctor should not ask financial information of the patient.
- In a **financial** context:
  - If a person makes a loan application, it would be **appropriate** to share financial information.

How can we avoid  
**unwarranted** or  
**indiscriminate** data  
collection?

- We should not collect data randomly.
- We should justify why certain pieces of information are collected:
  - **GDPR**: Data minimization
  - **GDPR**: Purpose should be specified

# iOS: Privacy Nutrition Labels



Have we given data subjects appropriate forms of **choice** in data sharing?

- Opt-in vs Opt-out Privacy settings



# Nike: Cookie Settings

*PS: They have now improved their cookie dialogue screen!!!*

STUDENTS NOW GET 10% OFF. [Learn More](#).

## YOUR COOKIE SETTINGS

Nike asks you to accept cookies for performance, social media and advertising purposes. Social media and advertising cookies of third parties are used to offer you social media functionalities and personalized ads. To get more information or amend your preferences, press the 'more information' button or visit "Cookie Settings" at the bottom of the website. To get more information about these cookies and the processing of your personal data, check our [Privacy & Cookie Policy](#). Do you accept these cookies and the processing of personal data involved?

[MORE INFORMATION](#) [YES, I ACCEPT](#)

You can always change your preference by visiting the "Cookie Settings" at the bottom of the page. View [Privacy & Cookie Policy](#) for full details.

## YOUR COOKIE SETTINGS

- Functional**

These cookies are required for basic site functionality and are therefore always enabled. These include cookies that allow you to be remembered as you explore the site within a single session or, if you request, from session to session. They help make the shopping cart and checkout process possible as well as assist in security issues and conforming to regulations.
- Performance**

These cookies allow us to improve the site's functionality by tracking usage on this website. In some cases these cookies improve the speed with which we can process your request, allow us to remember site preferences you've selected. De-selecting these cookies may result in poorly-tailored recommendations and slow site performance.
- Social Media and Advertising**

Social media cookies offer the possibility to connect you to your social networks and share content from our website through social media. Advertising cookies (of third parties) collect information to help better tailor advertising to your interests, both within and beyond Nike websites. In some cases, these cookies involve the processing of your personal data. For more information about this processing of personal data, check our [Privacy & Cookie Policy](#). De-selecting these cookies may result in seeing advertising that is not as relevant to you or you not being able to link effectively with Facebook, Twitter, or other social networks and/or not allowing you to share content on social media.

[DONE](#)

You can always change your preference by visiting the "Cookie Settings" at the bottom of the page. View [Privacy & Cookie Policy](#) for full details.

Are the terms of our data policy laid out in a **clear, direct**, and **understandable way**, and made **accessible** to all data subjects?

- Most **privacy policies** takes a legal perspective since they are written by lawyers.
- Privacy policies are not **updated frequently** to match data practices.

# Good Example

## Microsoft Privacy Statement

Last Updated: October 2021 [What's new?](#)

Your privacy is important to us. This privacy statement explains the personal data Microsoft processes, how Microsoft processes it, and for what purposes.

Microsoft offers a wide range of products, including server products used to help operate enterprises worldwide, devices you use in your home, software that students use at school, and services developers use to create and host what's next. References to Microsoft products in this statement include Microsoft services, websites, apps, software, servers, and devices.

Please read the product-specific details in this privacy statement, which provide additional relevant information. This statement applies to the interactions Microsoft has with you and the Microsoft products listed below, as well as other Microsoft products that display this statement.

Young people may prefer starting with the [Privacy for young people](#) page. That page highlights information that may be helpful for young people.

### Personal data we collect

How we use personal data

Reasons we share personal data

How to access and control your personal data

Cookies and similar technologies

Products provided by your organisation—notice to end users

Microsoft account

Collection of data from children

Other important privacy information 

### Product-specific details:

Enterprise and developer products 

Productivity and communications products 

Search, Microsoft Edge, and artificial intelligence 

Windows 

Entertainment and related services 

### Cookies

Most Microsoft sites use cookies, small text files placed on your device which web servers utilise in the domain that placed the cookie can retrieve later. We use cookies to store your preferences and settings, help with sign-in, provide targeted ads, and analyse site operations. For more information, see the [Cookies and similar technologies](#) section of this privacy statement.

### Personal data we collect

Microsoft collects data from you, through our interactions with you and through our products. You provide some of this data directly, and we get some of it by collecting data about your interactions, use, and experiences with our products. The data we collect depends on the context of your interactions with Microsoft and the choices you make, including your privacy settings and the products and features you use. We also obtain data about you from third parties.

If you represent an organisation, such as a business or school, that utilises Enterprise and Developer Products from Microsoft, please see the [Enterprise and developer products](#) section of this privacy statement to learn how we process your data. If you are an end user of a Microsoft product or a Microsoft account provided by your organisation, please see the [Products provided by your organisation](#) and the [Microsoft account](#) sections for more information.

You have choices when it comes to the technology you use and the data you share. When we ask you to provide personal data, you can decline. Many of our products require some personal data to provide you with a service. If you choose not to provide data required to provide you with a product or feature, you cannot use that product or feature. Likewise, where we need to collect personal data by law or to enter into or carry out a contract with you, and you do not provide the data, we will not be able to enter into the contract; or if this relates to an existing product you're using, we may have to suspend or cancel it. We will notify you if this is the case at the time. Where providing the data is optional, and you choose not to share personal data, features like personalisation that use such data will not work for you.

[Learn more](#)

[Top of page !\[\]\(7d13f531a752889c3606378b744d636d\_img.jpg\)](#)

### How we use personal data

Microsoft uses the data we collect to provide you with rich, interactive experiences. In particular, we use data to:



# Terms of Service, Didn't Read



**Facebook**

Grade E

- 🚫 Facebook stores your data whether you have an account or not.
- 🚫 The service can read your private messages
- 🚫 This service can view your browser history
- 🚫 Deleted content is not really deleted
- 👎 This service keeps user logs for an undefined period of time



**Startpage**

Grade A

👍 This service does not track you

👍 The service will resist legal requests for user information where reasonably possible

👍 IP addresses of website visitors are not tracked

👍 The cookies used by this service do not contain information that would personally identify you

👍 The service provides a complaint mechanism for the handling of personal data

View Details

Analyze Documents

Are data subjects  
being **appropriately  
compensated** for the  
benefits/value of their  
data?

- Compensation could be in the form of **money** (e.g., vouchers).
- Participation could also be **voluntary**.

# Participation Information Sheet (PIS)

## **What will happen if I decide to take part?**

### **Specify:**

- Kinds of data being collected (e.g. questions regarding X, Y or Z)
- Means of collection (e.g. questionnaire, interview, focus group)
- Duration of session

- If participant audio/video is being recorded
- How often, where, when

### **Compensation. [only required if applicable]**

You will be paid £X for your participation in this study [edit accordingly].

Have we considered what **control** or **rights** our data subjects should **retain** over their data?

- The users should be able to **withdraw**, **correct** or **update** their shared data.
- The data collector should make this **clear** in their privacy policies.

# Participation Information Sheet (PIS)

## **What are my data protection rights?**

The University of Edinburgh is a Data Controller for the information you provide [amend if student PI]. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights of correction, erasure and objection. [make restrictions clear if e.g. accessing data would affect outcome of study] For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit [www.ico.org.uk](http://www.ico.org.uk). Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at [dpo@ed.ac.uk](mailto:dpo@ed.ac.uk).

## 2. DATA STORAGE, SECURITY AND RESPONSIBLE DATA STEWARDSHIP

How can we  
**responsibly and safely**  
**store** personally  
identifying  
information?

- The data collectors should make it clear how they **store** their data.
  - Ethics Boards in organizations
  - Ethics Committees at universities

Have we **reflected on the ethical harms** that may be done by a data breach, both in the short-term and long-term, and to whom?

- Stakeholders can be **direct** or **indirect**.
- It is **very difficult** to **estimate the correct audience** who could be affected by a data breach.
- A **risk estimation** should be done in any case.

Have we made appropriate investments in our **data security/storage infrastructure** (relative to our context and the potential risks and harms)?

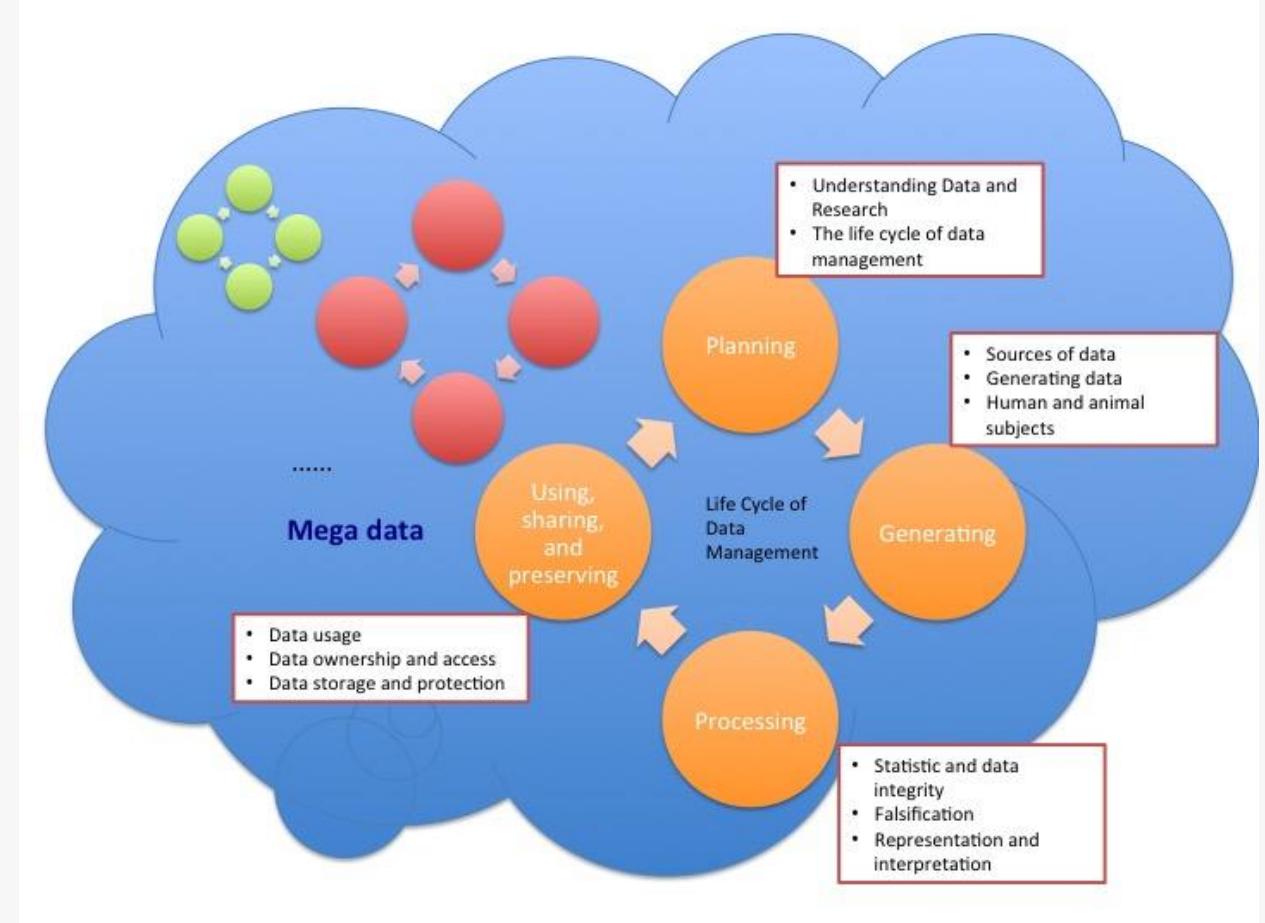
- A **good infrastructure** requires investment. Not all organizations can have it.
- The **context** in which data collection happens matters a lot.
  - sensitive vs non-sensitive

What **privacy-preserving techniques** do we rely upon, and what are their various advantages and limitations?

- Data pseudonymization
- Data anonymization
- Obfuscation
- Differential privacy

And so on...

Do we have an **end-to-end plan for the lifecycle of the data** we collect or use?



### 3. DATA HYGIENE AND DATA RELEVANCE

How **dirty** (inaccurate, inconsistent, incomplete, or unreliable) is our data, and how do we know?

- Dirty data means dirty models...
- Dirty models can do **significant harms** (e.g., criminal justice, hiring, finance etc.).

# Datasheets for datasets\*



- Bridging the gap between dataset creators and data consumers
- Good for:
  - Reproducibility
  - Increasing accountability and transparency
  - Mitigating unwanted biases
  - Deciding on the use of a dataset
- Similar datasets could be created based on datasheets

How long is this data likely to **remain** accurate, useful or relevant?

- For some applications, **time** aspect may be critical.
- Keeping historical or non-updated data may lead to **unethical decisions**.

# Use of historic criminal data

- Goal: **predicting criminal behavior** to allocate resources accordingly
  - for rehabilitation or for prison sentences in the US.
- AI assigns a **recidivism score** to people.
- Judges consider such scores in making decisions for defendants.

## AI is sending people to jail—and getting it wrong

Using historical data to train risk assessment tools could mean that machines are copying the mistakes of the past.

By Karen Hao

January 21, 2019



Low-income and minority communities have **historically** been disproportionately targeted by law enforcement

## 4. IDENTIFYING AND ADDRESSING ETHICALLY HARMFUL DATA BIAS

What inaccurate, unjustified, or otherwise harmful human biases are reflected in our data?

How might harmful or unwarranted bias in our data get magnified, transmitted, obscured, or perpetuated by our use of it?

- Biases can be explicit or implicit.
- Identification of harms is important, but the impact on the stakeholders should be thought about carefully.
- The wide use of non-documented datasets could perpetuate harmful/unwarranted bias.

## 5. VALIDATION AND TESTING OF DATA MODELS & ANALYTICS

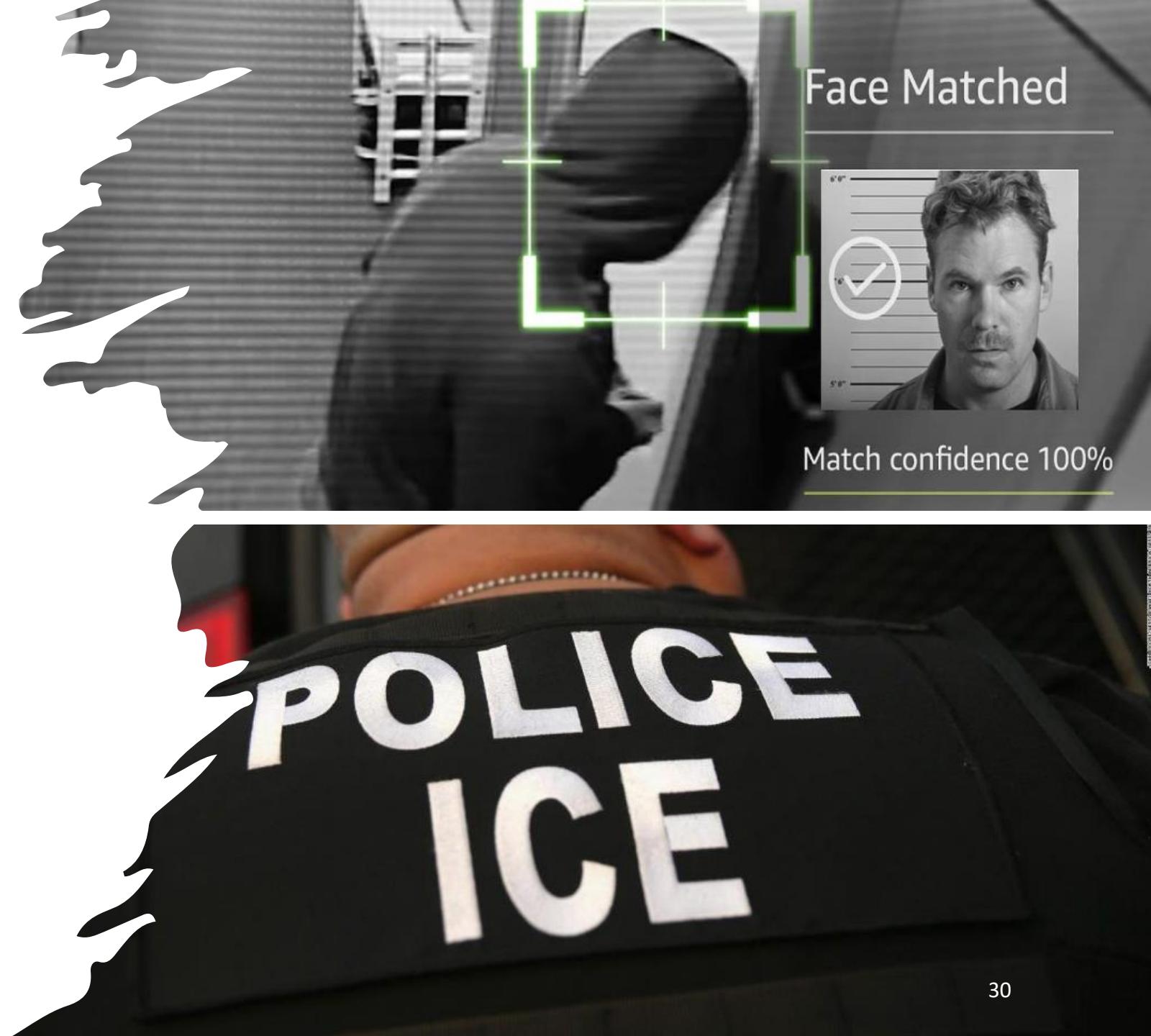
How can we ensure that we have **adequately tested** our analytics/data models to validate their performance?

Have we fully considered the **ethical harms** that may be caused by inadequate validation and testing?

- The **datasets** that we train/validate/test our models is critical.
- Validation/testing may be rushed because of **internal/external pressures**. This may lead to more ethical harms.

How can we test our data analytics and models to ensure their reliability across new, unexpected contexts?

What if the same system does harm to various stakeholders in another context?



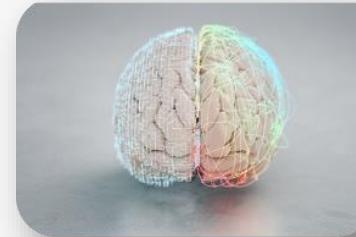
# 6. HUMAN ACCOUNTABILITY IN DATA PRACTICES AND SYSTEMS

Who will be designated as **responsible** for each aspect of ethical data practice?

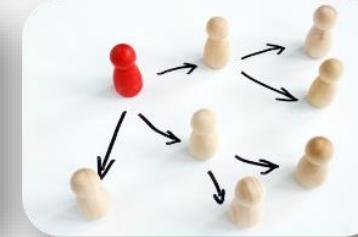
Who should and will be held **accountable** for various harms that might be caused by our data or data practice?



### Making Systems Answer: Dialogical Design as a Bridge for Responsibility Gaps in Autonomous Systems



What do philosophy & cognitive science teach us about responsibility gaps in AI?



Can moral theories of responsibility as answerability help us learn how to bridge these gaps?



Can we design AI systems to become more open to human answerability practices?



# 7.EFFECTIVE CUSTOMER/USER TRAINING IN USE OF DATA AND ANALYTICS

Have we placed data tools in appropriately skilled and responsible hands, with appropriate levels of **instruction** and **training**?

- Do organizations invest in **appropriate tools** to monitor how the data is used?
- Training staff is **essential**, since a user cannot do much without proper training.

# 8. UNDERSTANDING PERSONAL, SOCIAL, AND BUSINESS IMPACTS OF DATA PRACTICE

Have we fully considered how our data/data practice or system will be used, and how it might **impact** data subjects or other parties later on?

Has **sufficient input** been gathered from other stakeholders?

Has the **testing** of the practice taken into account how its impact might vary?

- **Diversity** in teams is crucial.
- **Focus groups** should be conducted to consult stakeholders of a system during the **design, development, and deployment** phases.



# Common Ethical Challenges

for Data Practitioners and Users

*\* based on Introduction to Data Ethics module (Part 2)*