

# PEAs in Pods

Community | Engagement | Research

PEAs in Pods  
Community | Engagement | Research



## EPSRC PEAs in Pods: Co-production Of Community Based Public Engagement For Data And AI Research

Funded by EPSRC Grant Ref: EP/W033488/1 2022



Engineering and  
Physical Sciences  
Research Council



Manchester  
Metropolitan  
University

# Activity 3.4 - Researching with communities in and beyond AI

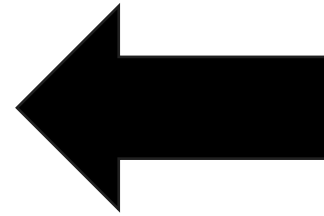


# Activity aims

- To understand the role of data in the AI Lifecycle
- To explore the role of communities in data collection and curation
- To discuss what are the Problems / Challenges / Benefits of citizen engagement



# What is the purpose of the engagement?



[This Photo](#) by Unknown Author  
is licensed under [CC BY-NC](#)

**STEM Ambassador**

**School Volunteer**

**UCAS talks**

**Science festivals**

**Career Talks**

# What is the purpose of the engagement?



*This Photo* by Unknown Author  
is licensed under [CC BY-NC](#)

**Citizen science**

**Co-creation**

**Co-production**

**Participatory AI**

# Communicating and **Understanding Artificial Intelligence**

**It's complicated!**

# Viewpoint – British Medical Journal

Source: British Medical Journal - Does “AI” stand for augmenting inequality in the era of covid-19 healthcare? | The BMJ

## Real world patterns of health inequality and discrimination



Unequal access and resource allocation

Discriminatory healthcare processes

Biased clinical decision making

## Discriminatory data



Sampling biases and lack of representative datasets

Patterns of bias and discrimination baked into data distributions

## Application injustices



Disregarding and deepening digital divides

Exacerbating global health inequality and rich-poor treatment gaps

Hazardous and discriminatory repurposing of biased AI systems

World → Data

Use ← Design

## Biased AI design and deployment practices

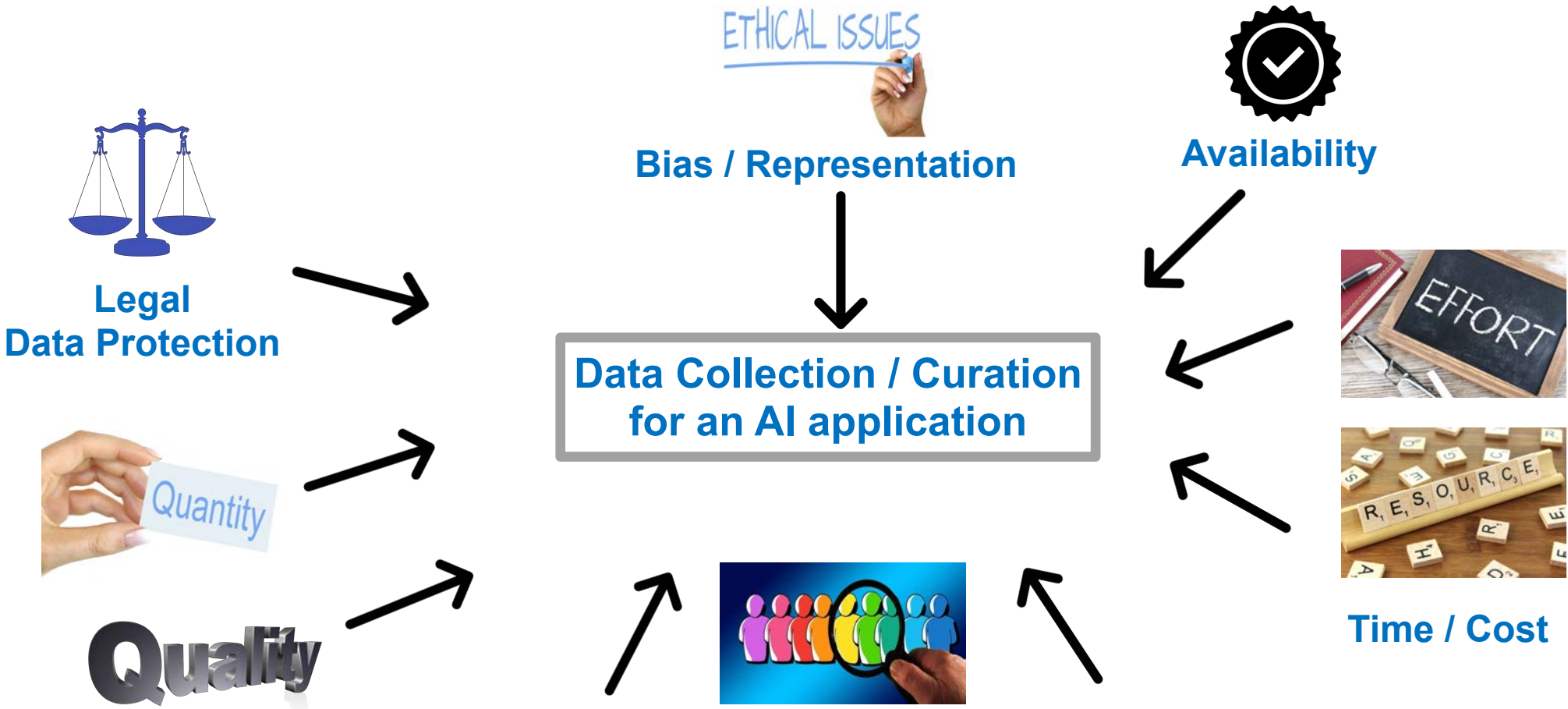


Power imbalances in agenda setting and problem formulation

Biased and exclusionary design, model building and testing practices

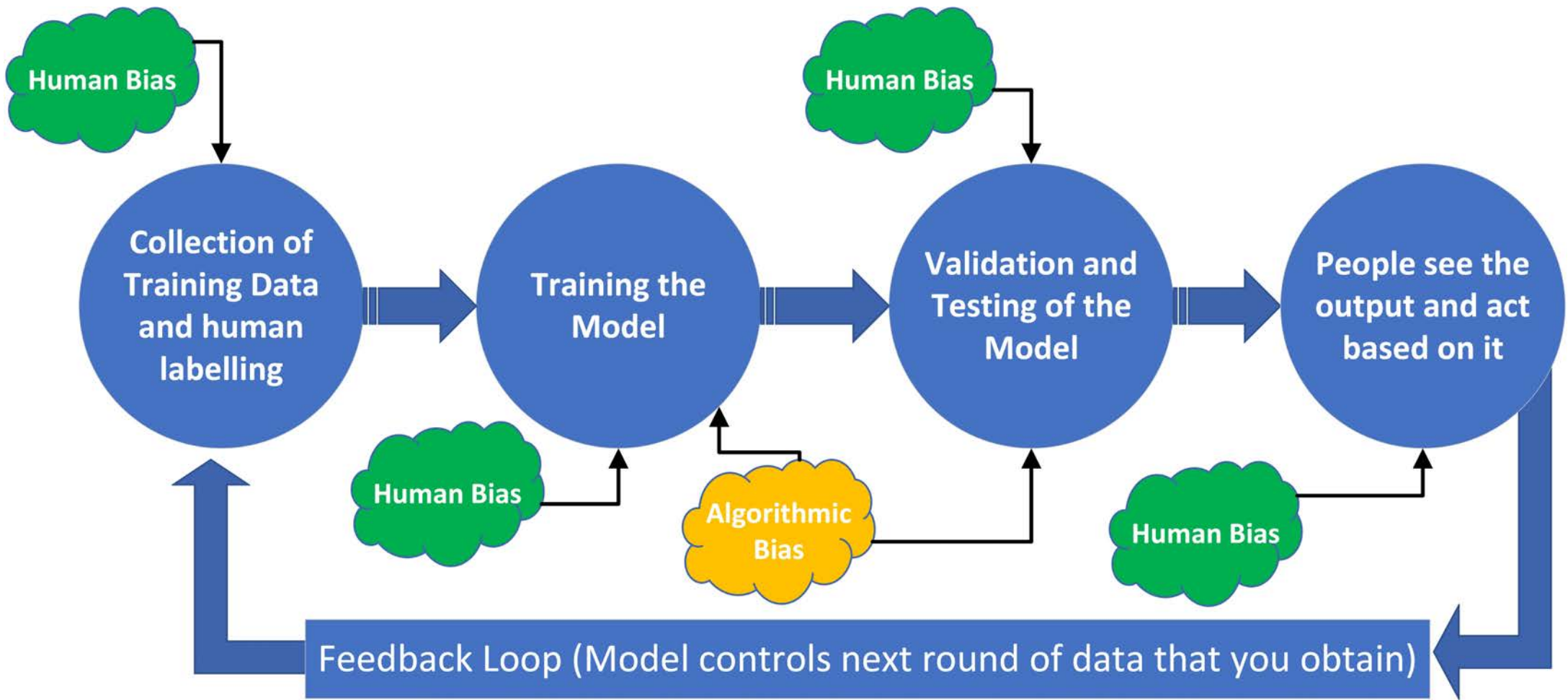
Biased deployment, explanation and system monitoring practices

# AI and Data - Challenges



**Methods: Harvesting / Web Scraping / Harvesting / Data Mining / Human / Automated / Crowdsourced / 3<sup>rd</sup> Party / In-house**

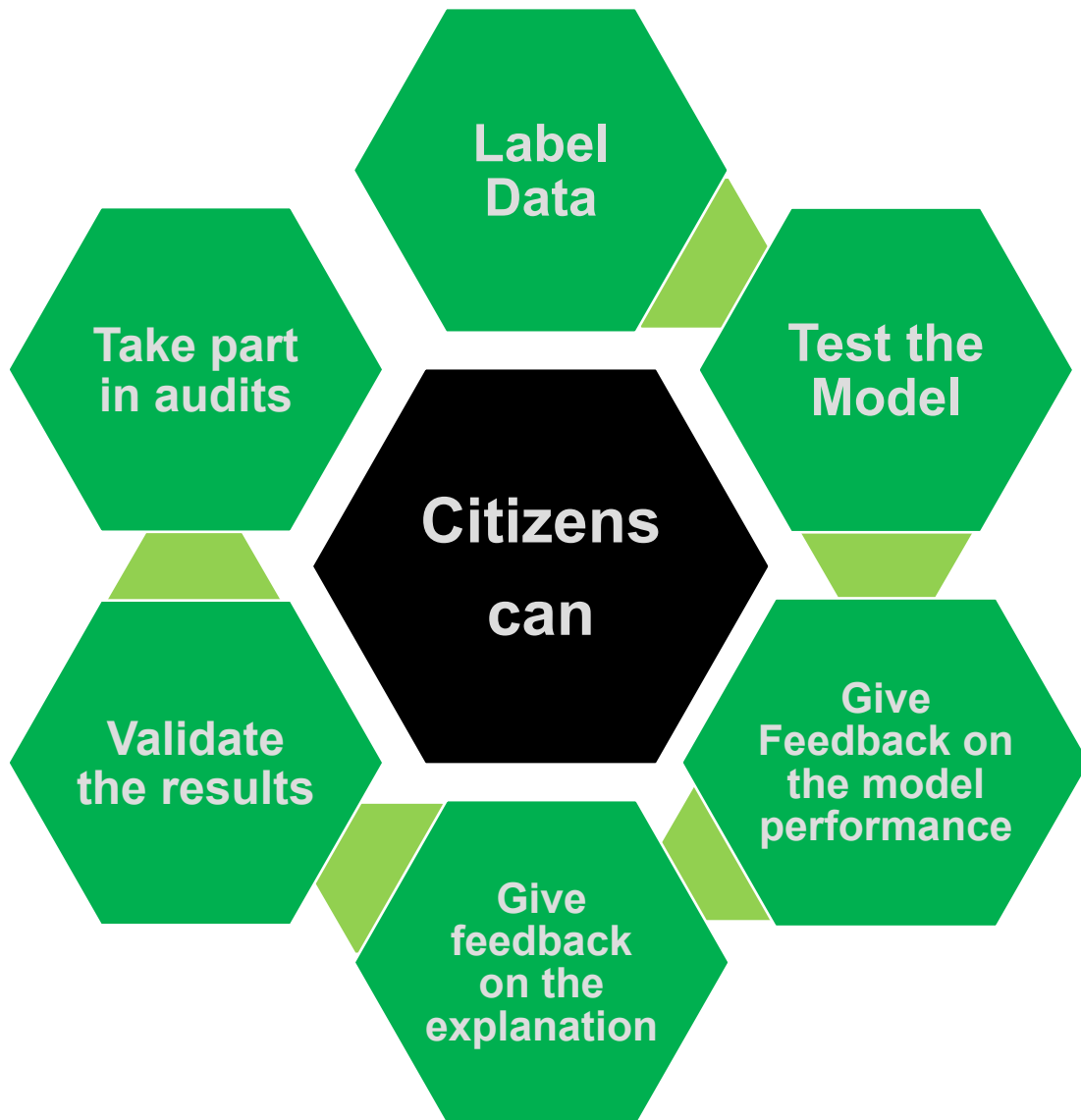
# Data in the AI Lifecycle





# What could be the role of Communities in Data Collection / Curation in the context of Artificial Intelligence?

# What are the Problems and Challenges and Benefits ?

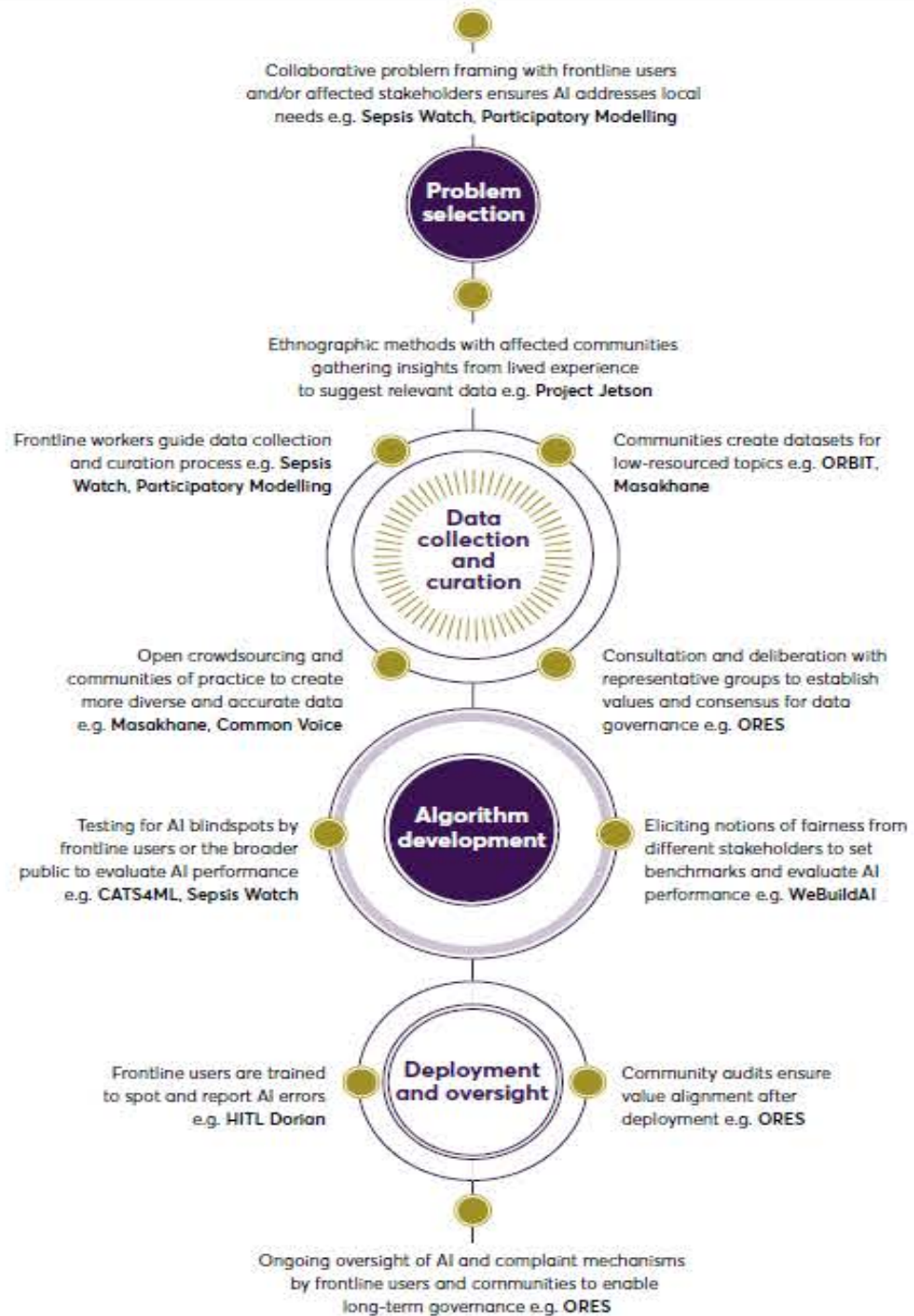


**Previous knowledge and Skills**

**Understanding of the application domain**

**How can it be done in practice?**

.....



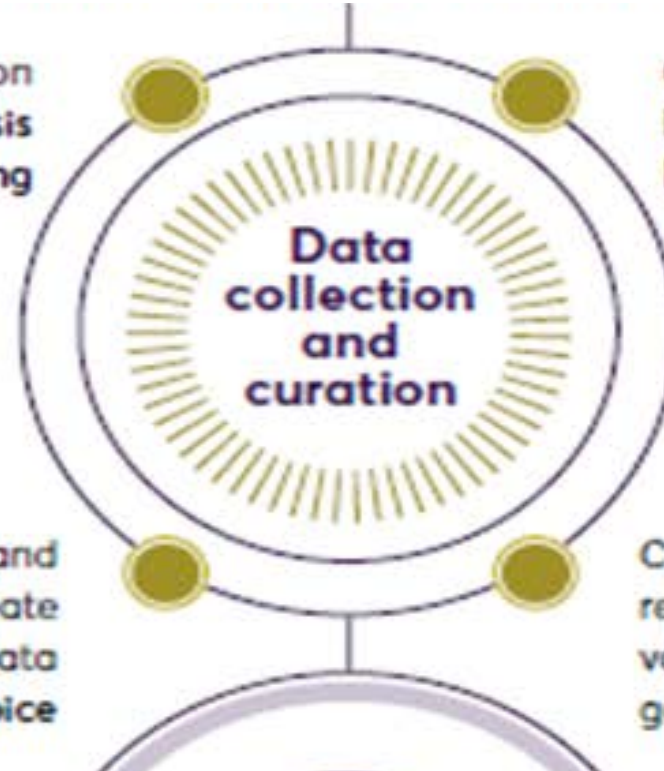
# Nesta - A framework for operationalising participatory AI

Source: [Participatory AI for humanitarian innovation \(nesta.org.uk\)](https://www.nesta.org.uk)

# Nesta - A framework for operationalising participatory AI

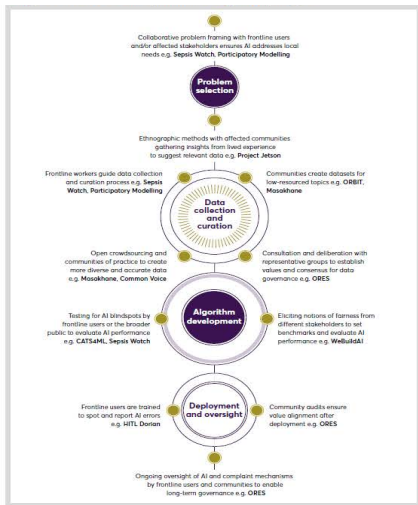
Frontline workers guide data collection and curation process e.g. Sepsis Watch, Participatory Modelling

Communities create datasets for low-resourced topics e.g. ORBIT, Masakhane



Open crowdsourcing and communities of practice to create more diverse and accurate data e.g. Masakhane, Common Voice

Consultation and deliberation with representative groups to establish values and consensus for data governance e.g. ORES



Source: Participatory AI for humanitarian innovation ([nesta.org.uk](https://www.nesta.org.uk))



**Does introduction of community voice into data collection / curation and labelling introduce more bias or better data representation? Why?**

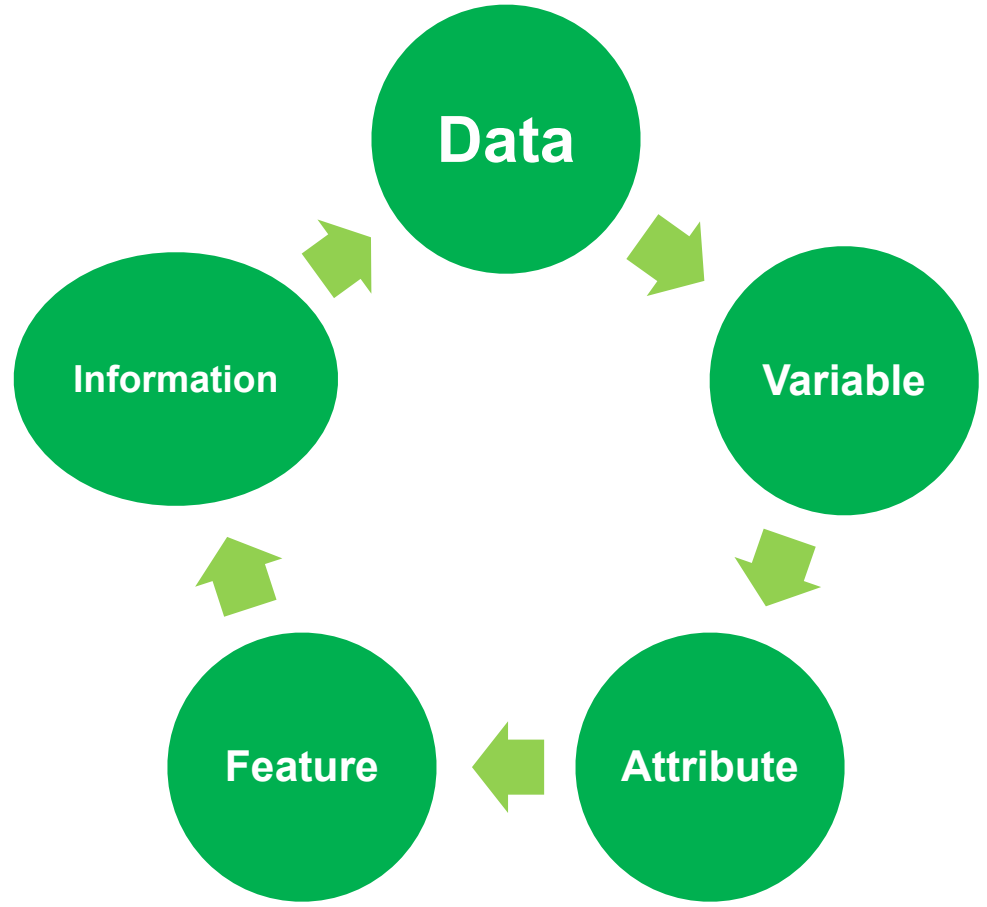
# CHALLENGES

## Communication of ideas

We may switch terminology in conversation depending on who we are talking to?

Conversations with communities need consistency

### Example - When describing how Artificial Intelligence “works” how do you refer to data?



# Quantitate Verses Qualitative Data

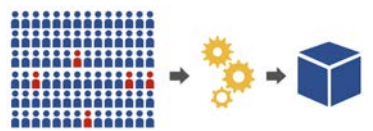
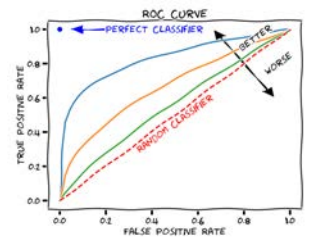
What data do we collect when undertaking machine learning?

Quantitative (Often)

- Accuracy.
- Confusion Matrix.
- Precision.
- Recall.
- F-Score.
- AUC(Area Under the Curve)-ROC .....

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)



What other data can we collect (examples)?

- EDI Monitoring data (Categorical e.g. gender)
- Surveys containing Likert Scale questions e.g. to access explainability of a model from a user perspective
- Free text answers → Quotes
- Consequence Scanning data to measure societal impact of new AI product and service
- Interview data (may require thematic coding)
- Specific Funder related themes e.g. trust, confidence

This Photo by Unknown Author is licensed under [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/)

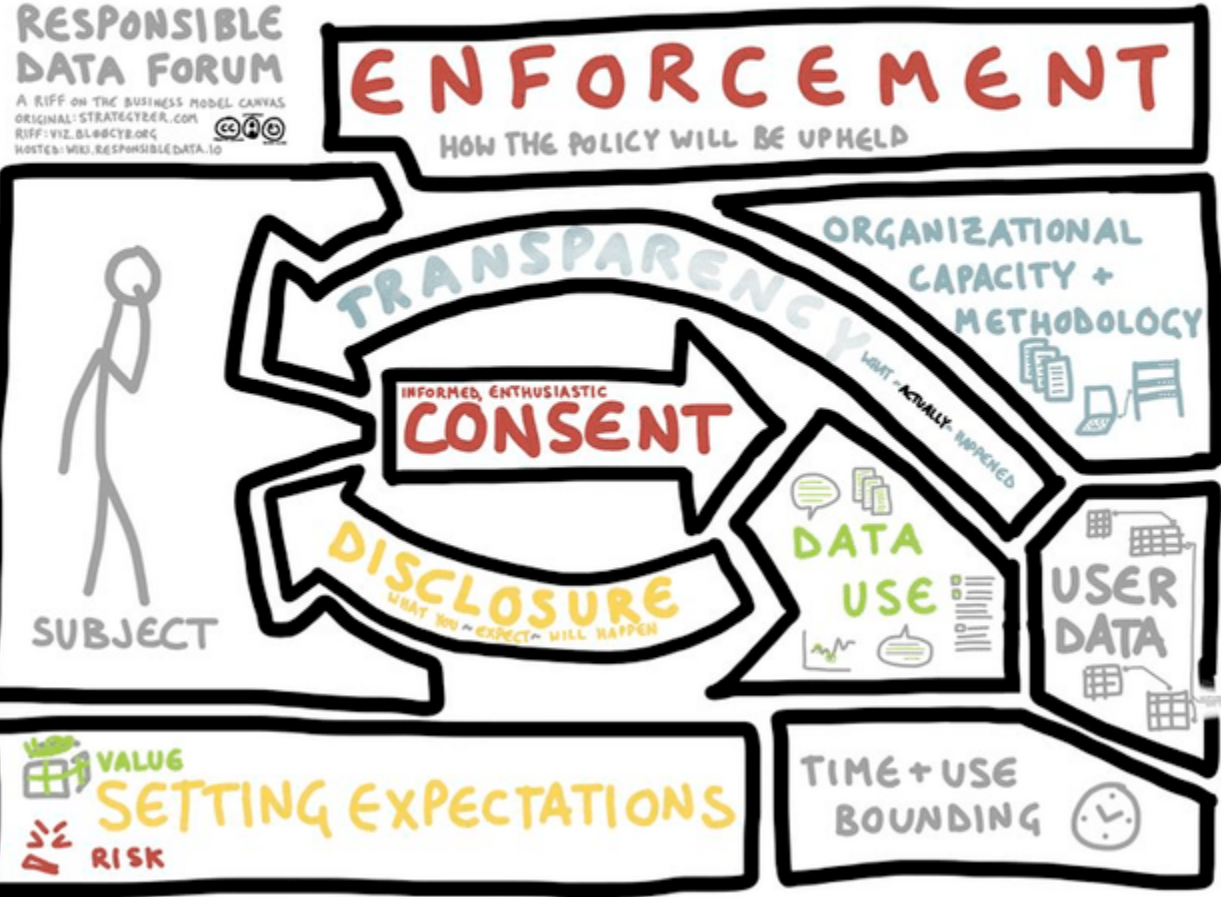
# Freely Given Informed Consent

“Information for participants should be provided in an accessible and comprehensive format, typically in written form (or in a form that participants can access after the end of the research interaction).

Time should be allowed for the participants to consider their choices and to discuss their decision with others if appropriate.”

Source: [Consent – UKRI](#)

# framework for CONSENT POLICIES



This Photo by Unknown Author is licensed under [CC BY-SA-NC](#)

# Participants Vulnerability

## UKRI (2025) Policy states researchers should consider

- participants' vulnerability
- potential negative consequences or lack of personal benefits from their involvement in research where these are expected
- providing appropriate information to elicit freely-given informed consent for participation as well as information regarding data deposit and data re-use (where deposit is possible)
- limits to confidentiality and occasions where this may occur
- legal requirements of working with the specific population (including Disclosure and Barring Service clearance)
- incentives and compensation for participation.



*This Photo* by Unknown Author is licensed under [CC BY](#)

Source: UKRI, 2025 <https://www.ukri.org/councils/esrc/guidance-for-applicants/research-ethics-guidance/research-with-potentially-vulnerable-people/>



**What are the Problems /  
Challenges / Benefits of citizen  
engagement in your own  
research?**