

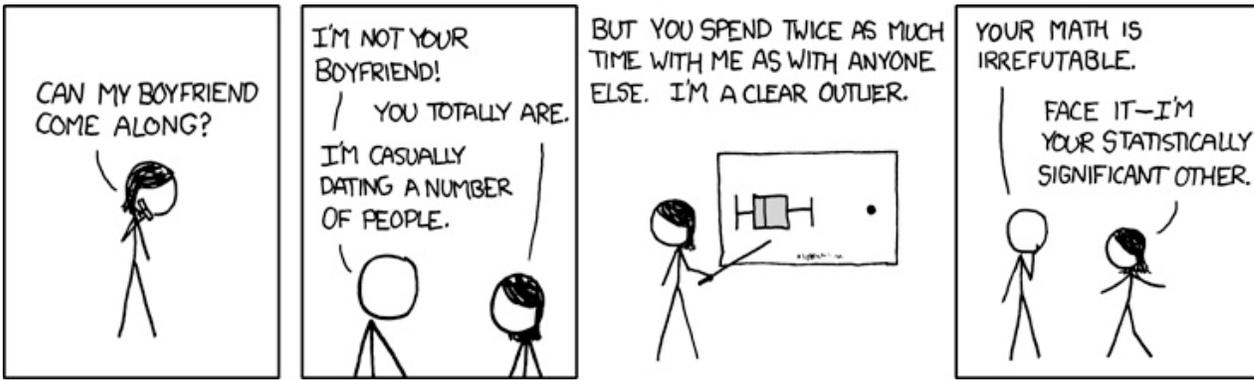
# Being Privately Open and Openly Private

a call for sustainable AI in the  
Archives

Yunhyong Kim  
([Yunhyong.Kim@Glasgow.ac.uk](mailto:Yunhyong.Kim@Glasgow.ac.uk))



University  
of Glasgow



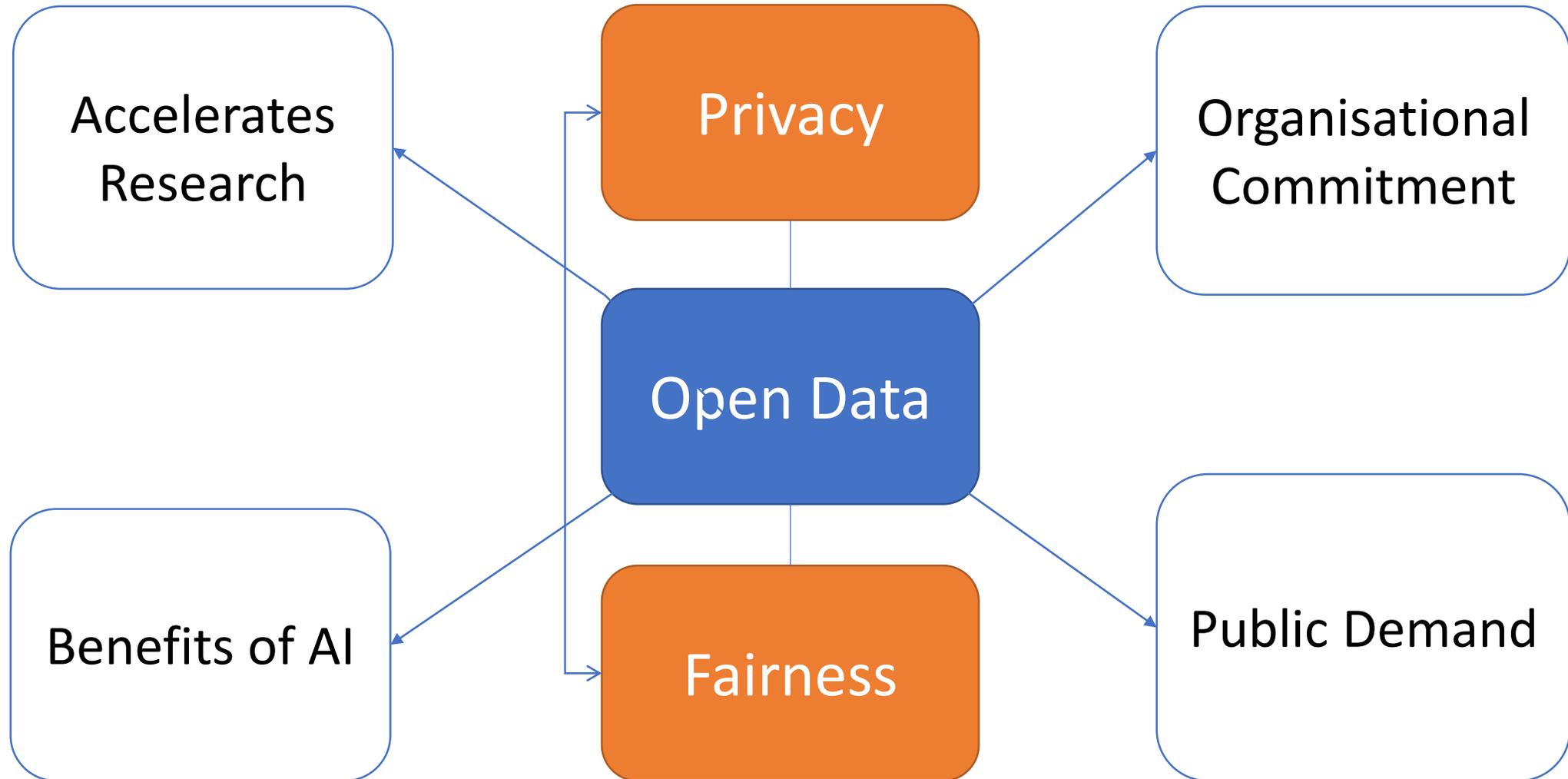
# Concerning Data

a conundrum: open data, privacy and fairness conflict?

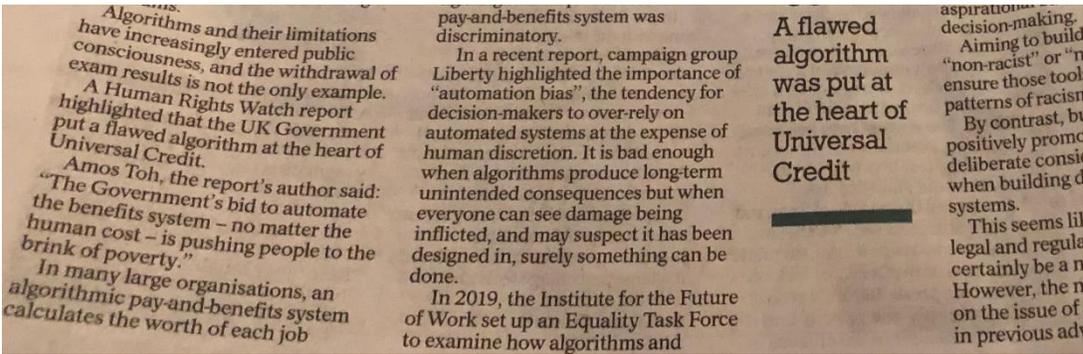
Image source: <https://xkcd.com/539/>

Creative Commons Attribution Non-commercial 2.5 <https://creativecommons.org/licenses/by-nc/2.5/>

# Open Data: a Common Goal



The Herald (Scotland): 14/11/2020



# Enter, AI Overlord

an opportunity to explore privacy and fairness?

DECISION MAKING

## Using Algorithms to Understand the Biases in Your Organization

by Jennifer M. Logg

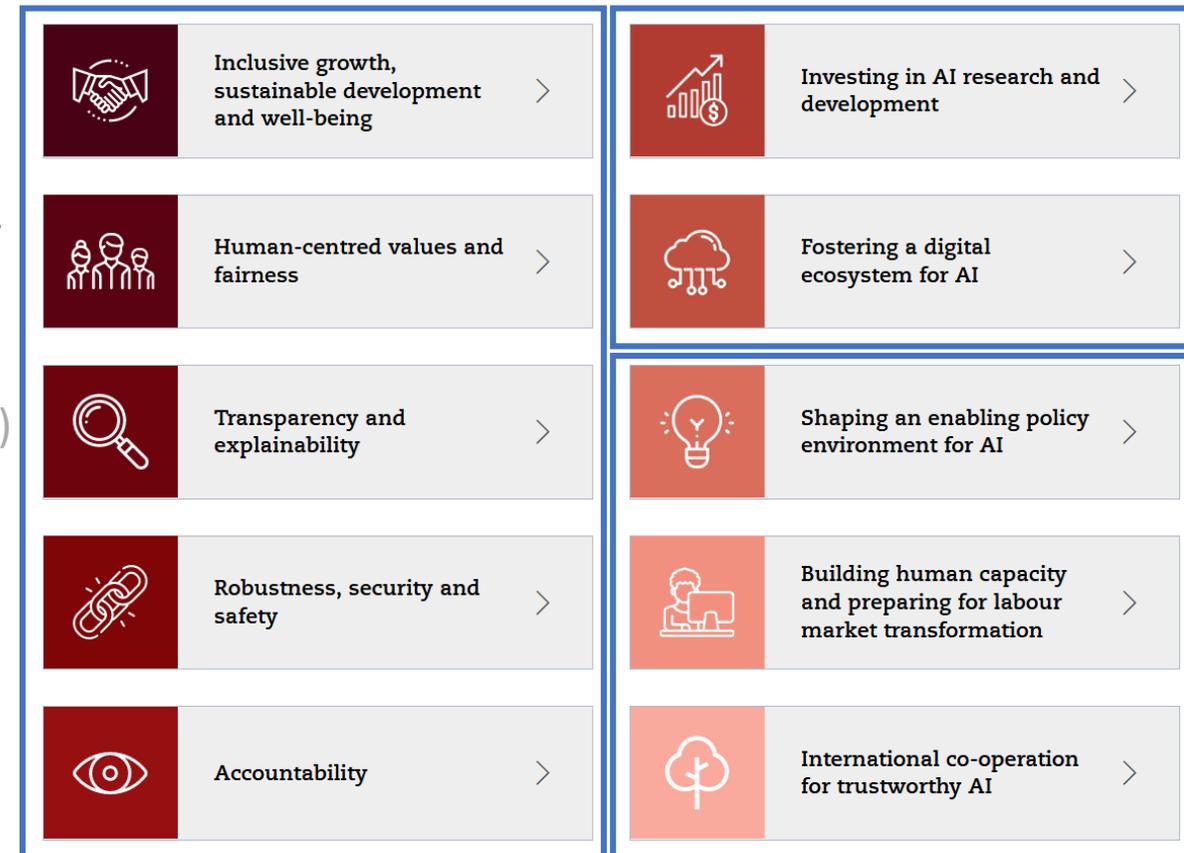
August 09, 2019

unearthing patterns that people have difficult detecting. When algorithms surface biases, companies should seize on this “failure” as an opportunity to learn when and how bias occurs. This way, they’re better equipped to debias their current practices and improve their overall decision making.

<https://hbr.org/2019/08/using-algorithms-to-understand-the-biases-in-your-organization>

# AI Spotlight on Privacy and Fairness

- Global Partnership on AI (GPAI)
  - International initiative to promote responsible AI use that respects human rights and democratic values
  - Conceived by Canada and France with 13 other founding members: Australia, the European Union, Germany, India, Italy, Japan, Korea, Mexico, New Zealand, Singapore, Slovenia, the United Kingdom and the United States.
  - Built on OECD AI Policy Observatory Principles
- Google AI Principles
- AI and Machine Learning for Coders (Laurence Moroney) - 2020
  - Chapter 20. AI Ethics, Fairness and Privacy
- Building an Anonymization Pipeline (Luk Arbuckle, Khaled El Emam) - 2020
- Building Machine learning pipelines (Hannes Hapke; Catherine Nelson)
  - Chapter 14. Data Privacy for Machine Learning
- The AI Ladder (Rob Thomas, Paul Zikopoulos) – 2020
  - Chapter 3. How to Overcome AI failures and Challenges



# AI Spotlight on Privacy and Fairness

- Global Partnership on AI (GPAI)
  - International initiative to promote responsible AI use that respects human rights and democratic values
  - Conceived by Canada and France with 13 other founding members: Australia, the European Union, Germany, India, Italy, Japan, Korea, Mexico, New Zealand, Singapore, Slovenia, the United Kingdom and the United States.
  - Built on OECD AI Policy Observatory Principles
- Google AI Principles
- AI and Machine Learning for Coders (Laurence Moroney) - 2020
  - Chapter 20. AI Ethics, Fairness and Privacy
- Building an Anonymization Pipeline (Luk Arbuckle, Khaled El Emam) - 2020
- Building Machine learning pipelines (Hannes Hapke; Catherine Nelson)
  - Chapter 14. Data Privacy for Machine Learning
- The AI Ladder (Rob Thomas, Paul Zikopoulos) – 2020
  - Chapter 3. How to Overcome AI failures and Challenges

- Google AI principles:
  - Socially beneficial
  - Avoid creating or reinforcing unfair bias
  - Built and tested for safety
  - Be accountable to people
  - Incorporate privacy design principles
  - High standards of scientific excellence
  - Be made available for uses that accord with these principles
- Fairness
  - Is ML actually necessary
  - Design and implement metrics from day one
  - Build a minimum viable model and iterate
  - Infrastructure that supports rapid redeployment
- Explainability: What-If , Facets tools
- Privacy
  - Federated learning: data never leaves your device, model shared

# AI Spotlight on Privacy and Fairness

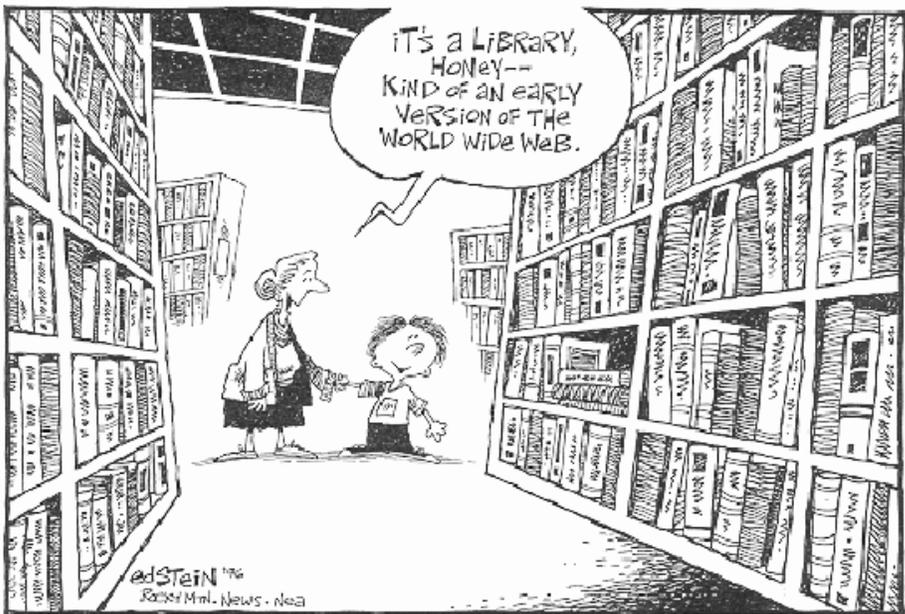
- Global Partnership on AI (GPAI)
  - International initiative to promote responsible AI use that respects human rights and democratic values
  - Conceived by Canada and France with 13 other founding members: Australia, the European Union, Germany, India, Italy, Japan, Korea, Mexico, New Zealand, Singapore, Slovenia, the United Kingdom and the United States.
  - Built on OECD AI Policy Observatory Principles
- Google AI Principles
- AI and Machine Learning for Coders (Laurence Moroney) - 2020
  - Chapter 20. AI Ethics, Fairness and Privacy
- Building an Anonymization Pipeline (Luk Arbuckle, Khaled El Emam) - 2020
- Building Machine learning pipelines (Hannes Hapke; Catherine Nelson)
  - Chapter 14. Data Privacy for Machine Learning
- The AI Ladder (Rob Thomas, Paul Zikopoulos) – 2020
  - Chapter 3. How to Overcome AI failures and Challenges

- Privacy and AI: could cause reduced accuracy and cost in time
- Privacy and fairness: privacy (e.g. deletion of personal information) could affect the possibility of gauging bias of data
- Privacy and open data: closed parts of the data could reduce the utility of the data
- Complexity of privacy: identifiability is not just about direct collection of personal data
- Data privacy: key questions
  - Who are you trying to keep the data private from
  - Which parts of the system can be private and which can be exposed to the world
  - Who are the trusted parties that can view the data
- Control of data passed to creators
- Encrypted machine learning
- Scrubbing the data for example using regex and named entity recognition
- Differential privacy: adding noise
- Federated learning: only sharing models
- Private Aggregation of Teacher Ensembles

# Failures and Opportunities

- Global Partnership on AI (GPAI)
  - International initiative to promote responsible AI use that respects human rights and democratic values
  - Conceived by Canada and France with 13 other founding members: Australia, the European Union, Germany, India, Italy, Japan, Korea, Mexico, New Zealand, Singapore, Slovenia, the United Kingdom and the United States.
  - Built on OECD AI Policy Observatory Principles
- Google AI Principles
- AI and Machine Learning for Coders (Laurence Moroney) - 2020
  - Chapter 20. AI Ethics, Fairness and Privacy
- Building an Anonymization Pipeline (Luk Arbuckle, Khaled El Emam) - 2020
- Building Machine learning pipelines (Hannes Hapke; Catherine Nelson)
  - Chapter 14. Data Privacy for Machine Learning
- The AI Ladder (Rob Thomas, Paul Zikopoulos) – 2020
  - Chapter 3. How to Overcome AI failures and Challenges

- Data challenges
  - Data silo
  - Lack of data
  - Disorganised data
  - Data quality
  - FAIR principle (findable, accessible, interoperable, reusable)
- Cultural Challenge
  - Potential employees with knowledge, skillset and experience are rare
  - Company culture and organisational silos
- Building trust
  - Fully traceable provenance
  - Lineage of the model and training data
  - Inputs and outputs of AI recommendation
  - Explainability

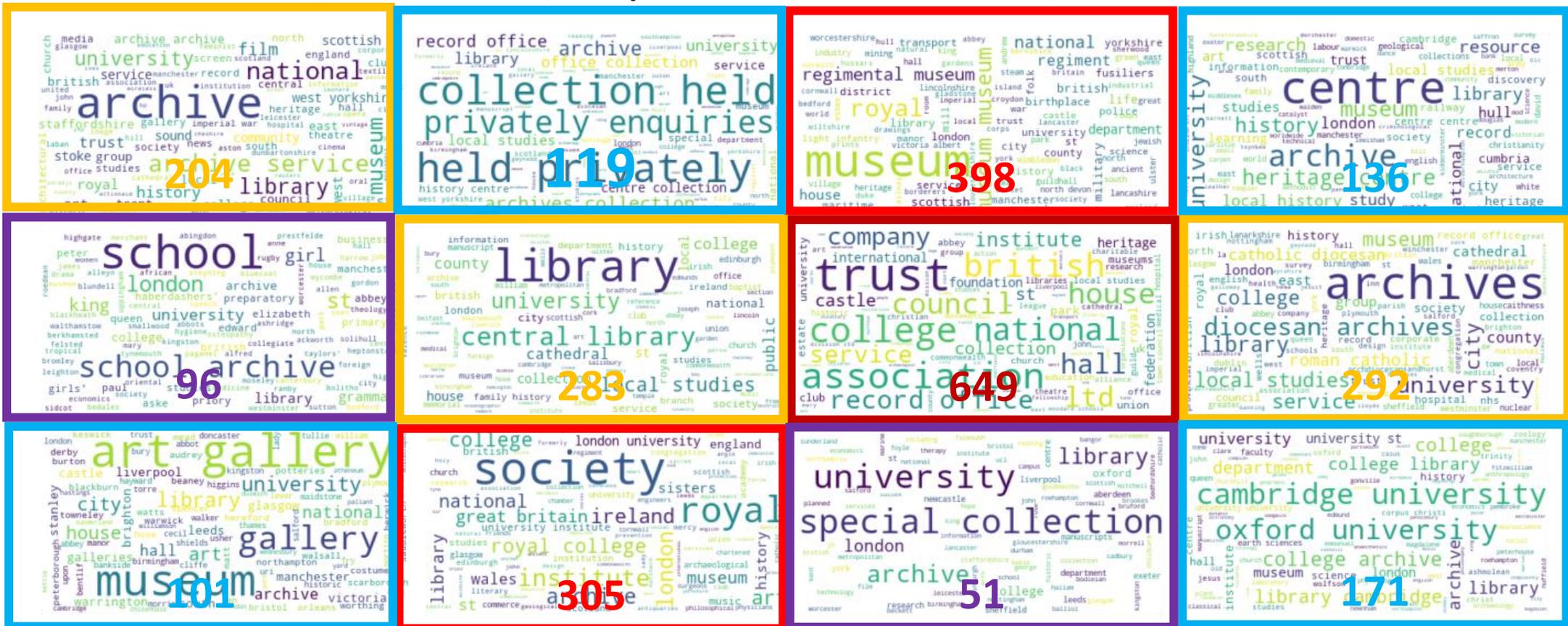


# Archival Challenges

A gap in the debate about AI and privacy

# Archives in the UK & RI Clustered

2805 listed at [discovery.nationalarchives.gov.uk](http://discovery.nationalarchives.gov.uk) – England 2309, Scotland 287, Wales 92, Republic of Ireland 88, Northern Ireland 29.



# Data Challenge

- Data held in the archives are growing both in volume and complexity. Archive data suffer from data challenges (e.g. silo, lack of data, quality).
- Most solutions for privacy from the AI point of view tends to assume data as a something like a statistical database (data collected with a confidentiality pledge - statistics that, by their production, do not compromise the privacy of those who provided the data). Much of archive data sets are not statistical databases
- In many cases, the public and researchers request to see the raw data. In the arts and humanities many researchers combine close and distant reading for analyses – discombobulating approaches like differential privacy.
- Even where the sensitive nature of documents are understood, still onsite access adopted as a solution

# Cultural Challenge

- Current areas of concern:
  - AI to make sense of data and boost research and engagement
    - Search, retrieval, recommendation, visualisation
  - AI as an approach for facilitating archival process
    - Metadata extraction, selection and appraisal, sensitivity review
- Neglected:
  - Establishing archive data as an indispensable source for AI research
  - Fostering archive culture to become part of the AI community
  - Developing the trusted environment for AI opportunities to emerge painlessly but safely

# Being Privately Open and Openly Private

The principles and examples

# A call for sustainable AI in the Archives

- Developing Open Archive Data for AI: transforming data (e.g. differential privacy) to
  - Share with the AI community and researchers
  - Share with other archives
  - e.g. Kaggle; Open ML public/private datasets; Jupyter Notebooks
- Developing Open AI for Archive Data
  - Building infrastructure for sharing AI models across archives (e.g. federated learning).
  - Establishing methodology for learning from failures across archives
- Agreed environment in place for trusted AI related data and model sharing

# Thank you! Questions?

