Practicum meeting: How to govern data and enable trustworthy AI

Meeting of 07/01/2022

Quick recap

Working on these points

- Federated Trust
- Distributed Trust
- Data catalog management
- How to implement trust?
- Finding a research question
- Testing/Audit of the proposed model (spoiler: very hard…)

Outcomes of reviewing current scientific literature about the topic

What can we implement?

- Fairness
- Transparency
- Auditability
- Privacy
- Responsibility

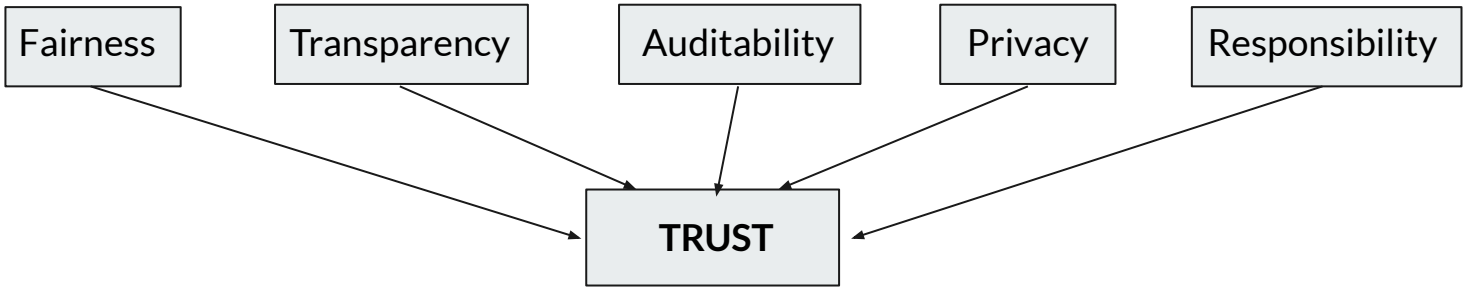
What can't we implement?

- Actual international organization approval e.g. WHO
- Simulate a vast amount of different devices connected i.e. heterogeneous environment
- simulate high volume of traffic

[1]-[4]

How to enable trust?

There is not a strict definition of Trust in data governance, hence researchers tend to establish some rules that can theoretically lead to enabling trust. [1]-[4]



Fundamental properties to enable TRUST

Fairness

Goals

- Reduce BIAS
- Improve accuracy of predictions
- Enable ethical and legal requirements for data governance

How to implement it

- Certification of dataset (completeness, consistency, accuracy, etc..)
- Requires human interaction (certification process)
- Definition of roles and responsibilities
- In machine learning tasks, may require other organizations to execute the same task to double-check and/or combine different results

[2],[3]

Transparency

Goals

- Make decision making processes clear (trace of decision steps)
- Improve awareness of stakeholders
- Support the outcomes of AI with vast collection of data (sometimes it is difficult to accept a decision)

How to implement it

- Explainable AI (We are not experts.. TO STUDY)
- Binding data with AI outcomes with possibility of step by step rollback (traceability)
- Collection of documents for the use of data (data subjects will know where their data are used and why, with possibility of enabling anonymization)
- Creation of standard of explainability of AI (scale or levels of explainability)

[1],[3]

Auditability

Goals

- Create trust by auditing the correctness of other organizations working on the same dataset
- Create a public and shared **consensus** in a “democratic” way (if many organizations review the results, it is probably correct)
- Promote the use of shared knowledge

How to implement it

The idea would be emulate somehow what happens in academic context.

Reports on data and AI outcomes are generated and peer-reviewed by other organizations.

Producing peer-reviewed knowledge generates points, if cited or used by other organizations the authors gain other points, the review process can be done in a form of “lottery” where some random organization are chosen (with equal probability) and asked to review a report, if they don’t do it in a certain time other randomly chosen organization can do it (and gain points).

This points can be used to create a sort of H-index of authoritativeness

Privacy

Goals

- Be compliant with law and ethics requirements e.g. GDPR
- Gather and share information without harming any individual
- Promote and support awareness of data usage, importance and also privacy risks

How to implement it

- Federated learning (not so much explicable)
- Robust anonymization protocol for personal data, especially sensitive personal data
- Possibility to withdraw immediately consent to use personal data

[1]-[3]

Responsibility

Goals

- Attribute property of data to authors
- Attribute property of AI to authors
- Trust is a human concept, it must be enabled by human responsibility

How to implement it

- Licensing system for data, reports and AIs
- Attribution of names and surnames of authors of AIs and Data Collectors.
- Handover tracking system

Before we continue

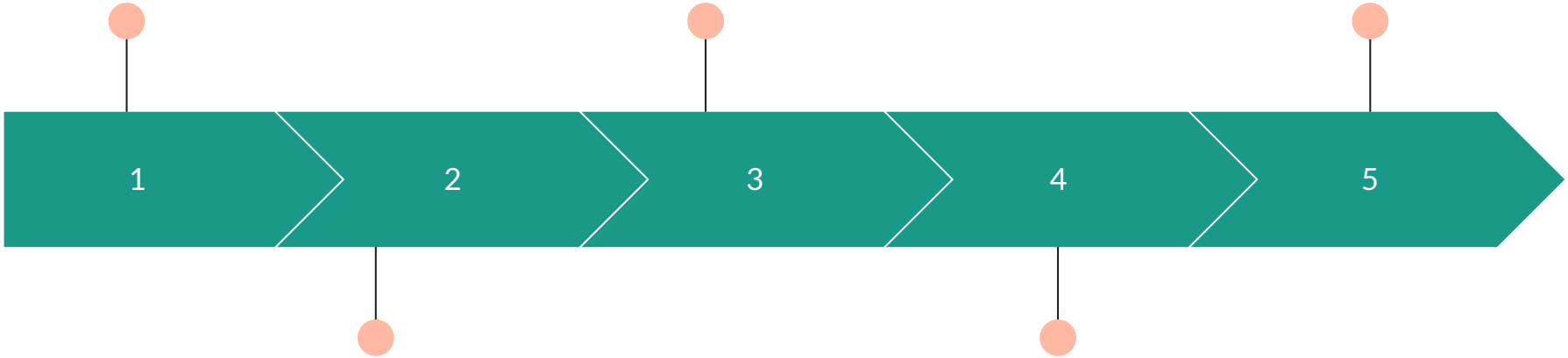
- Could this project be interesting in the academic context?
- Are these proposals enough? We want to do a very good job!
- Are these proposals too much? Limited time!
- Suggestions?
- Many papers about this topic propose a model but they don't really implement it, can we do the same (hybrid theoretical analysis for some parts and actual software implementation for others)?
- Lack of experience in data governance, we might need your help

How can we effectively define a robust research question?

Scientific Literature review

Construct a data model

Implementation
Test
Future work



Define technical & functional requirements

Learn environment requirements & specs e.g. CKAN

Homework

Any final comment / suggestion would be appreciated.

Notes:

Thank you

References

[1] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50–60, May 2020, ISSN: 1558-0792. DOI: 10.1109/msp.2020.2975749. [Online]. Available: http://dx.doi.org/10.1109/MSP.2020.2975749.

[2] E. Bertino, A. Kundu, and Z. Sura, “Data transparency with blockchain and ai ethics,” J. Data and Information Quality, vol. 11, no. 4, Aug. 2019, ISSN: 1936-1955. DOI: 10.1145/3312750. [Online]. Available: https://doi.org/10.1145/3312750.

[3] S. Reddy, S. Allan, S. Coghlan, and P. Cooper, “A governance model for the application of ai in health care,” Journal of the American Medical Informatics Association, vol. 27, Nov. 2019. DOI: 10.1093/jamia/ocz192.

[4] M. Janssen, P. Brous, E. Estevez, L. S. Barbosa, and T. Janowski, “Data governance: Organizing data for trustworthy artificial intelligence,” Government Information Quarterly, vol. 37, no. 3, p. 101 493, 2020, ISSN: 0740-624X. DOI: https://doi.org/10.1016/j.giq.2020.101493. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0740624X20302719