

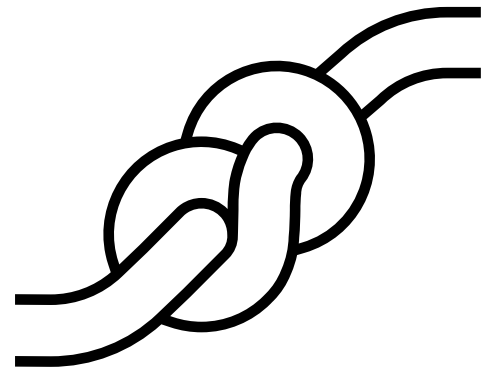
Data governance and Record-keeping in the Artificial Intelligence Act

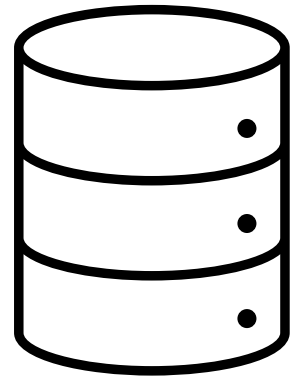
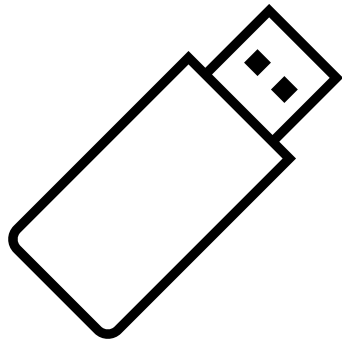
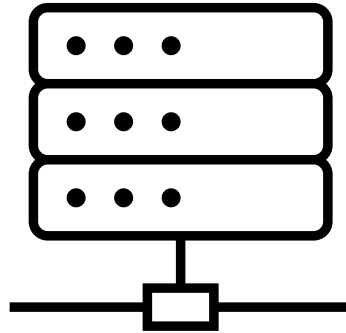
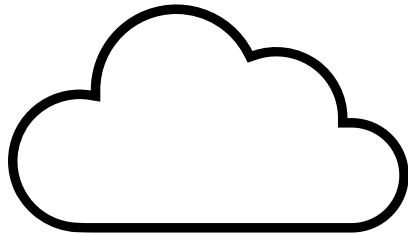
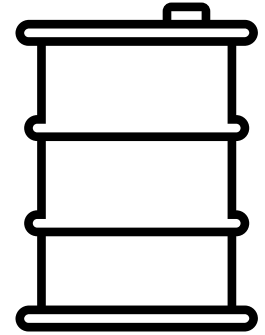
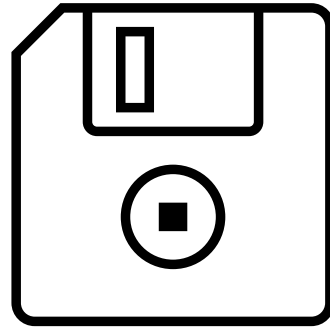
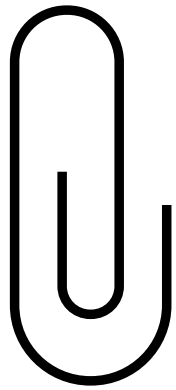
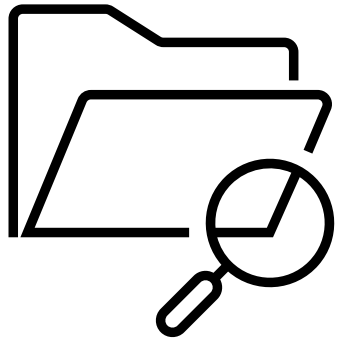
ANAHÍ CASADESÚS DE MINGO

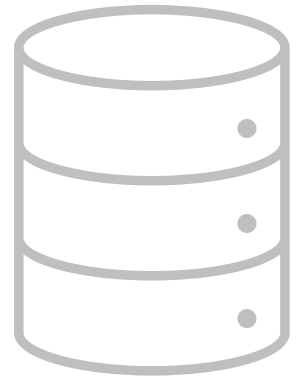
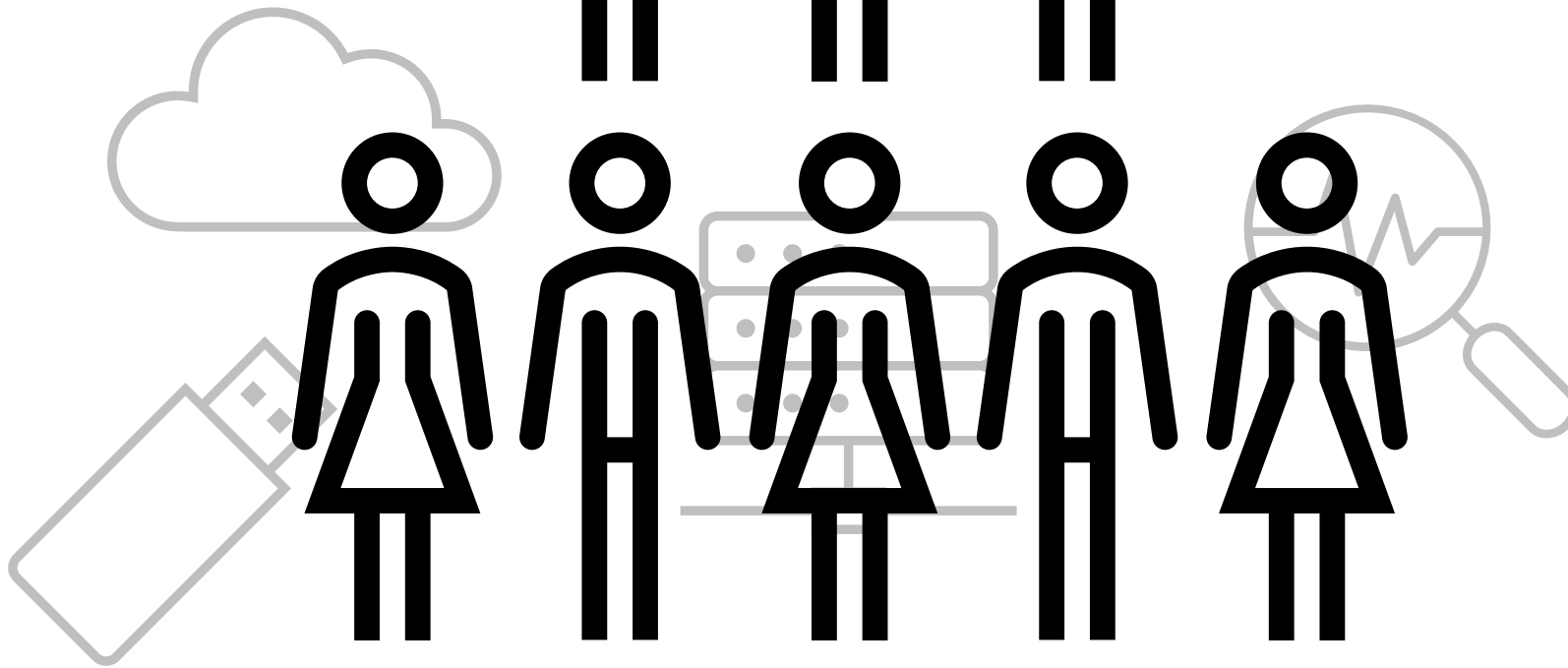
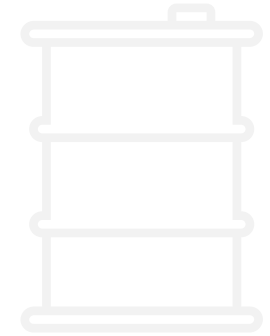
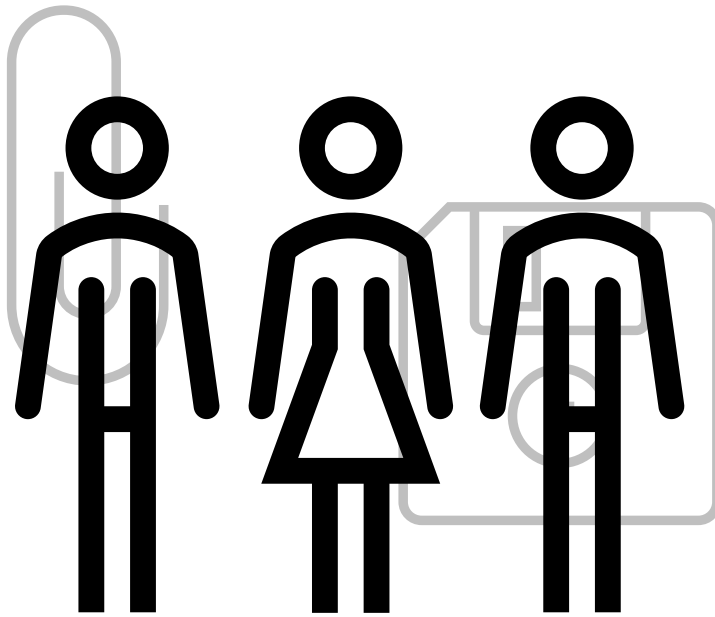
19/09/2023

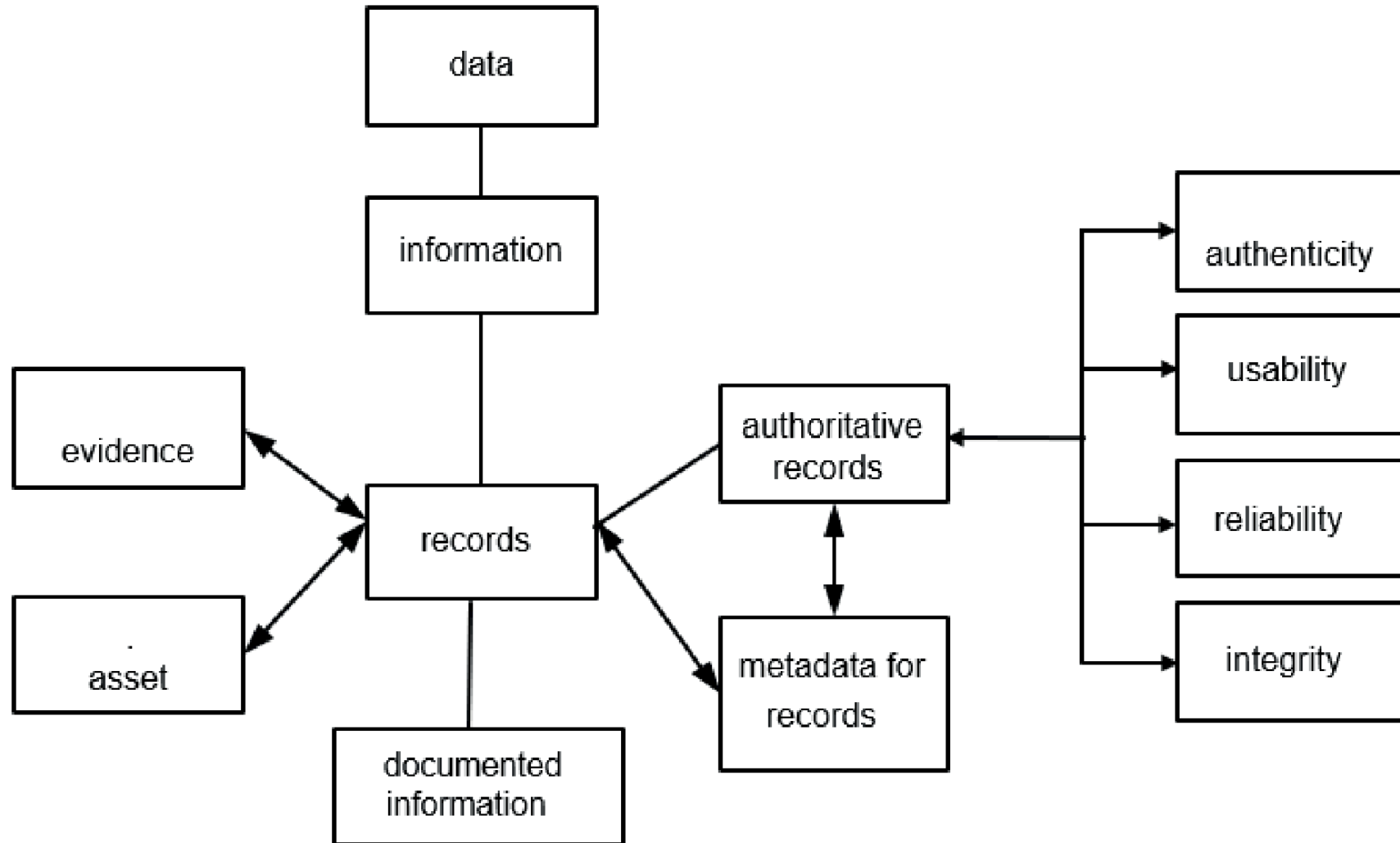
Data Governance ↔ Record-keeping

1. Conceptual framework
2. Data Quality
3. Data Governance in the AIA
4. Record-keeping in the AIA
5. Discussion









1. Data Governance ↔ Record-keeping

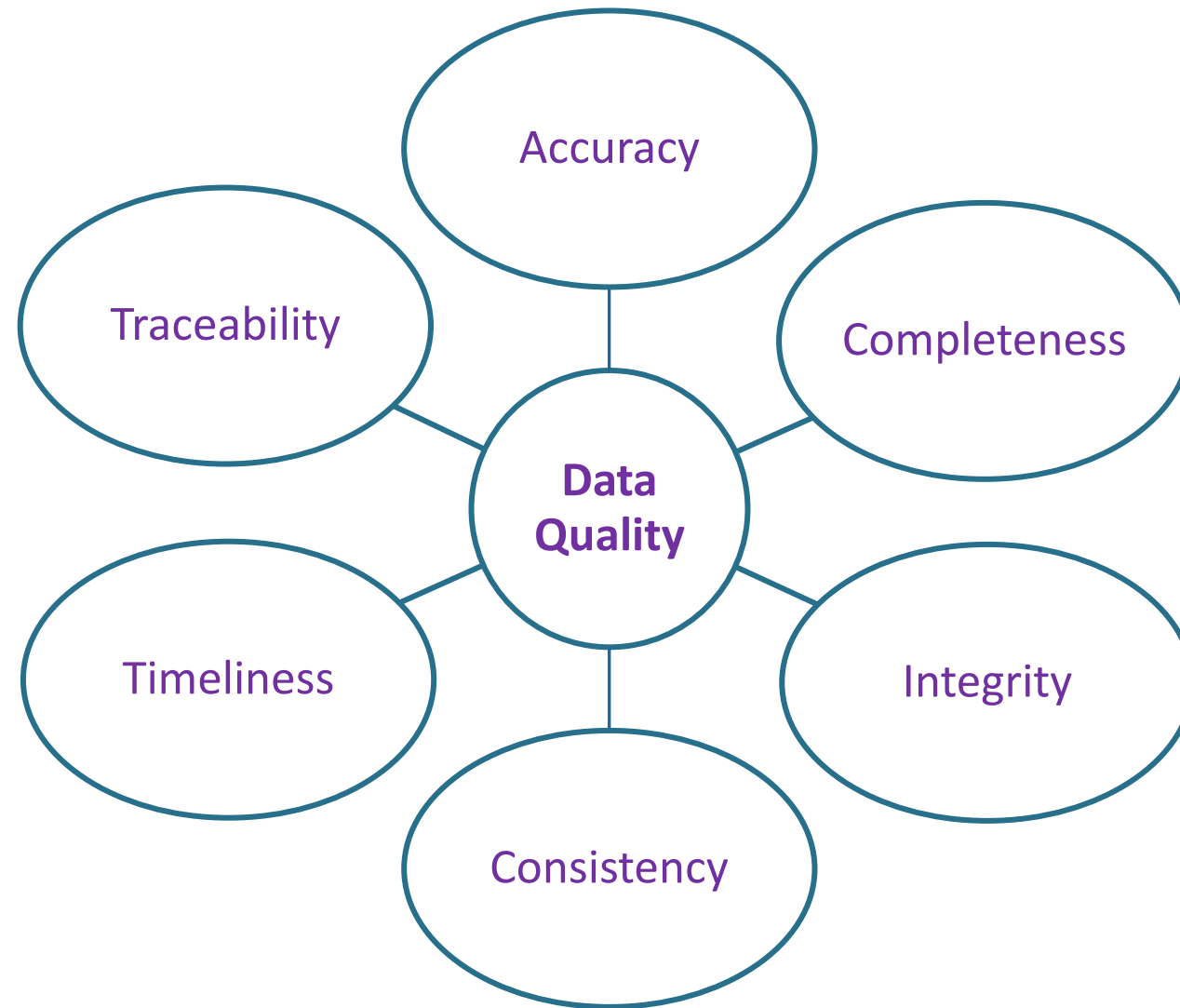
- Systems working **as a one**
 - Interoperability
 - Data quality
- Technology is not a magic solution
- **Interdisciplinary** approach



2. Data Quality

The quality of the data sets first depends on the quality of the data itself

- Nothing in the AIA
- Nothing in the Data Governance Act (EU)
- Not one model of quality dimensions (a lot of proposals)
- All agree on: requirements for the **use** for which it was intended (**context**, user)



Quality data dimensions, following Zhang

3. Data Governance in the AIA

- One article → Article 10
- Focused on high-risk AI systems
- Training of models with data – data training sets
- Refers to **quality** criteria, but not specifies any
- List of **good practices** for data governance

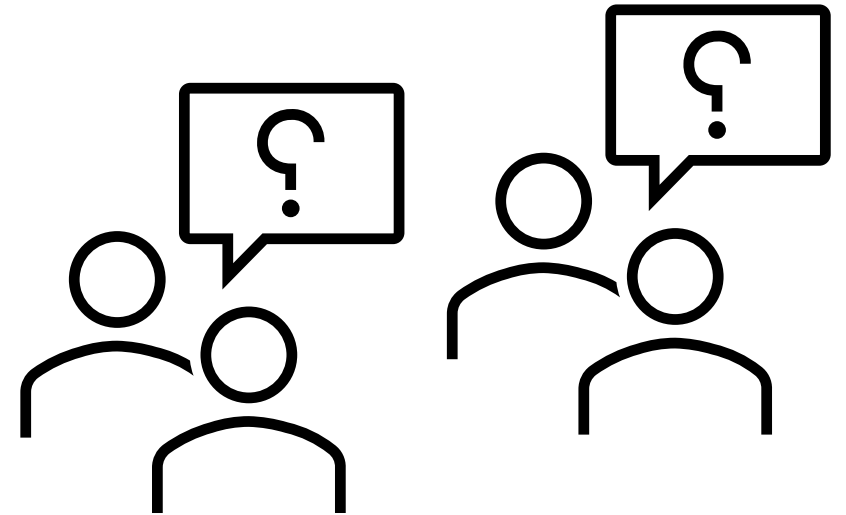


3. Data Governance in the AIA

- List of good practices for data governance in AIA
 - Data collection
 - Data preparation processing operations
 - Availability, quantity, suitability
 - Possible biases
 - Gaps or shortcomings
 - ...

3. Data Governance in the AIA

- No mention to:
 - Data governance models
 - Data quality dimensions
 - Acceptable / Not acceptable
 - Non-high-risk AI systems



4. Record-keeping in the AIA

Any set of information can be managed as a record

- One article → Article 12
- **Automatic** recording of events while high-risk system is operating
- Main objective = **Traceability** of the system's functioning throughout its life cycle

4. Record-keeping in the AIA

- Traceability of the system's functioning throughout its life cycle
 - Situation that may result in the AI system presenting a risk(art. 65.1)
 - Substantial modification
 - Post-marked monitoring (art. 61)

4. Record-keeping in the AIA

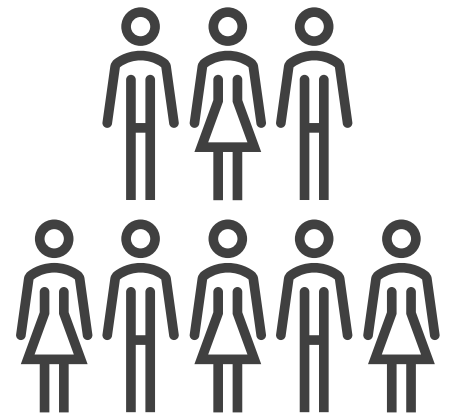
- Traceability of the system's functioning throughout its life cycle (annex III)
 - Period of each use
 - Reference database against which input has been checked
 - Input data
 - Identification of natural persons involved in the verification of results (art. 14.5)

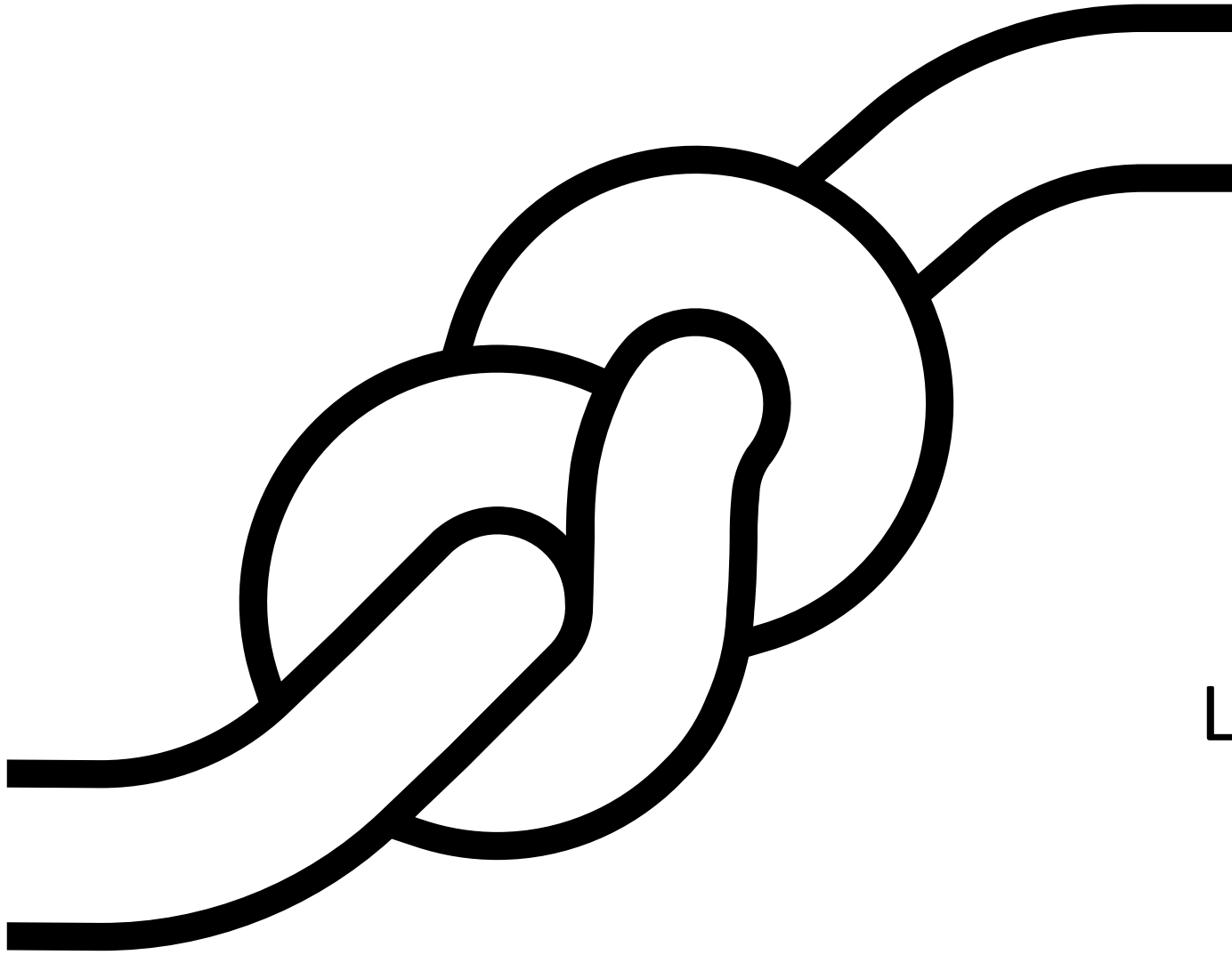
4. Record-keeping in the AIA

- No mention to the characteristics of records systems (applicable to AI systems?)
 - Reliable
 - Secure
 - Compliant
 - Comprehensive
 - Systematic

5. Discussion

- Only reference to high-risk AI systems
- No clear criteria about data quality
- What to do once bias is in the system?
- How to *get rid of* bias in existing data sets?
- Information governance is not only a technological solution





Let's work together !