TERRITORI APERTI

Enabling Trustworthy Predictions using ML and Multi-Modal Data

Andrea Bianchi, Giordano D'Aloisio, Antinisca Di Marco, Giovanni Stilo

University of L'Aquila

UNIVERSITÀ DEGLI STUDI DELL'AQUILA

COMUNE DELL'AQUILA

SOBIGDATA.it ITALIAN RESEARCH INFRASTRUCTURE

Finanziato dall'Unione europea NextGenerationEU

PARTNER

CNR-ISTI

USRA Ufficio Speciale per la Ricostruzione dell'Aquila

UNIVERSITA' DEGLI STUDI D'ANNUNZIO

MUR

CGIL CISL UIL

Problem

- Multi-modal data (i.e., data that span different types and contexts) can help stakeholders in many domains
- The complexity of multi-modal data leads to challenges in creating models able to integrate the knowledge derived from each data type
- In addition, to be used by different groups and populations ML systems must be *trustworthy* (i.e., explainable, fair, and private)

Case study: Medical Scenario

- Regulations on ethical, privacy and security issues [1][2][3] on data
- Growing volume and variety of medical data are collected by different entities. This means data belonging to different context, with different meanings, in different physical places
- Data are generated in multiple ways (images, genetics, natural language reports, electronic records) to cover different contexts
- Becoming medical specialist means having lot of experience. For AI algorithms experience comes from large, varied and high-quality dataset! (difficult to find in healthcare)

1. https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html
2. https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html
3. https://edps.europa.eu/data-protection/our-work/subjects/health_en

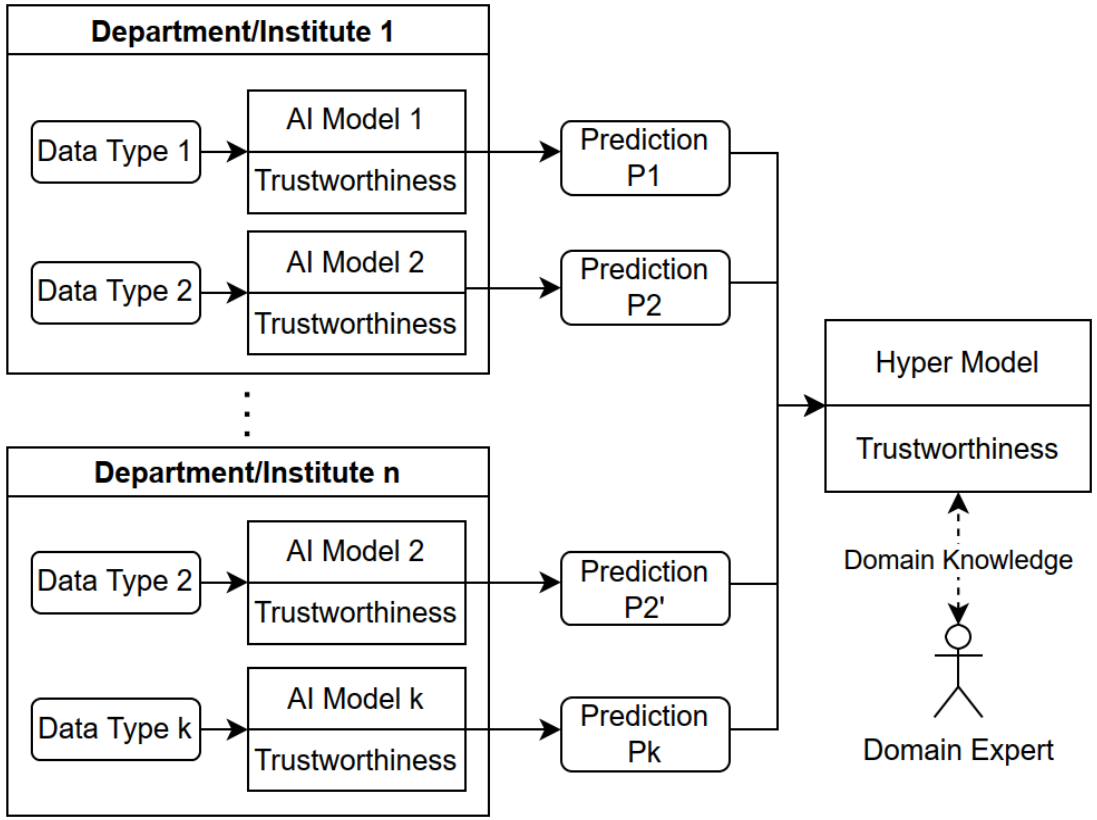
Challenges

In this work, we aim to solve the following challenges:

1. *How can we analyze multi-modal data in order to learn specific (i.e., sub-domain) and integrated (i.e., domain) knowledge?*
2. *How can we assure the trustworthiness of the predictions obtained using multi-modal data, both for specific and integrated knowledge?*

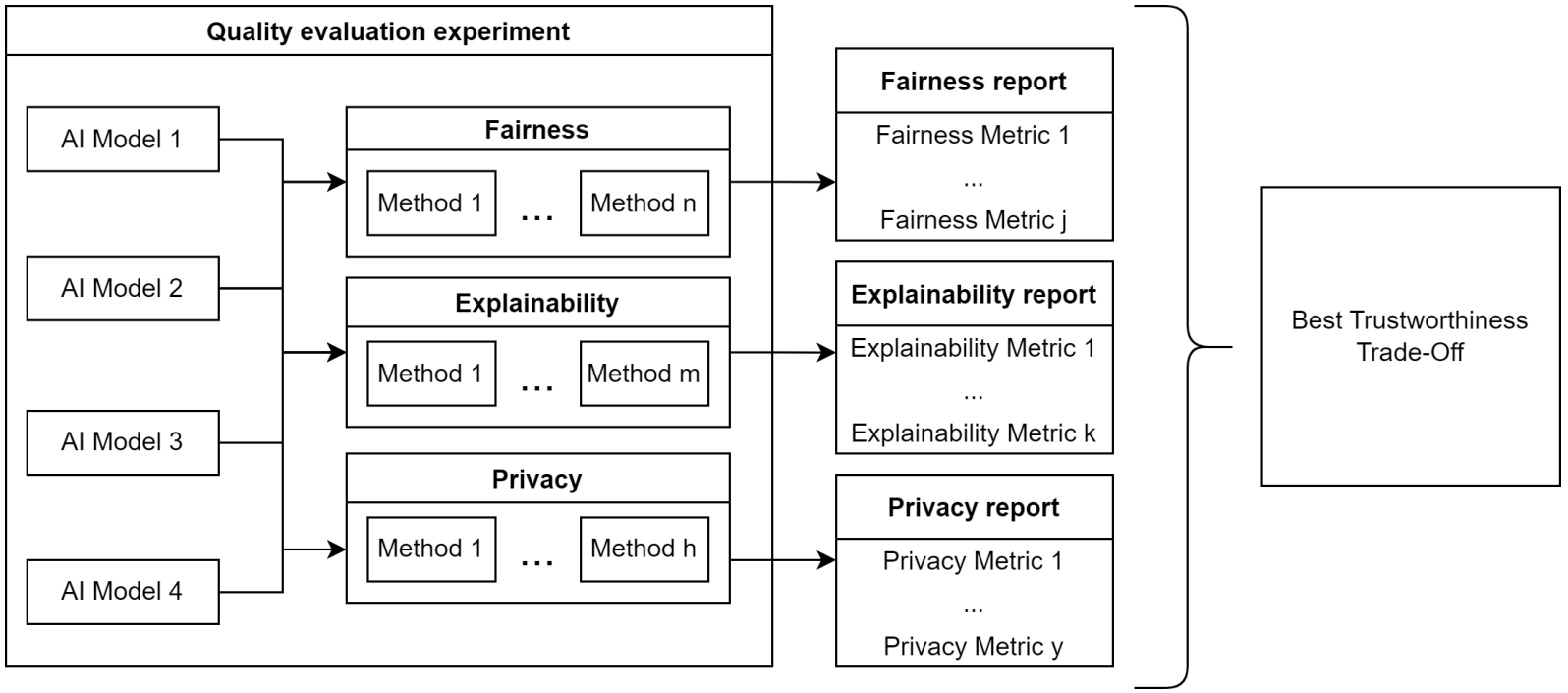
The need for a trust “global” learning

- Medical domain requires to build trust of clinicians and patients (quite difficult)
- Different models to analyze different data types (ex, CNN for images, Language Models for text, …)
- The predictions are sent to a new model, which will be driven by these predictions
- Knowledge to semantically link the predictions obtained by the different models (creating semantics)



Trustworthiness Assurance Process

- Evaluate each AI Model with each trustworthy method and compute related metrics
- We model this process as a Software Product Line formalized by Extended FM



Final remarks

- It is very difficult to find complete and high-quality datasets in the medical domain. Very often we encounter small datasets, semantically unrelated and biased
- On such dataset the creation of small learning pipelines could be better in maintaining the trustworthy in the predictions
- This approach can be generalized and extended to other domains characterized by the presence of multimodal data, which cannot be shared and where the quality requirement of trustworthiness is required