HI Methodologies and Metrics Challenge

RQ1: What is HI?

RQ2: How to evaluate the HI-ness of HI systems?

RQ3: How to systematically design HI systems?

AIM A: Developing guidelines that will lead to standards for researchers and practitioners to design and assess HI systems

AIM B: Developing tailored metrics that can accurately measure relevant aspects of hybrid intelligence, including people-related metrics, and metrics for process and progress

AIM C: Developing tools that can help researchers towards a systematic investigation of HI systems

AIM D: Discriminating Hybrid Intelligence from other related fields, such as General Artificial Intelligence or Human-AI collaboration

Cluster 1
Validation and Extension of the Quality Model

Delivered:
- Organized 3 HI competitions; Collected and published new dataset of 48 HI scenarios; Submitted joint paper to JAAMAS

Plans
- Finalize joint paper
- Extend quality model with additional quality attributes and metrics, informed by clusters 2 and 3

Cluster 2
Human-Focused Metrics (performance, role of MM, human cognition)

Plans:
Creation of a categorization on frequently used terms related to subjectivity, e.g., personalization, disagreement, etc.
- Creation of typology of dimensions for subjective input (ambiguity vs. concordance)
- Creation of desiderata for subjective LLMs and evaluating accordingly by mapping existing metrics to the respective criteria and proposing new metrics

Cluster 3
Human-Focused Metrics (performance, role of MM, human cognition)

Plans:
Plans
Creation of a document of metrics used for assessing tasks related to the scientific assistant (i.e. measuring novelty, shared understanding, usefulness, ...). Later, mapping them to CARE principles

Cluster 5
General guidelines for HI teams

Delivered:
Organized 1 workshop at the consortium meeting in April '25; Identified shortcomings of current guidelines and desiderata for future ones
Currently; Exploration of empirical literature to identify pitfalls in human-AI collaboration

Plans
Creation of new guidelines based on collected desiderata while addressing potential pitfalls; Joint paper

Cluster 6
Engineering effort: Online collaboration platform for Defining and Understanding Hybrid Intelligence

Delivered:
Submitted proposal

Plans:
Develop platform

Cluster 8
From high level to low level properties of HI teams

Plans:
Systematic mapping
Joint paper