



FAIR @ CircleU - Human-centered AI



Funded by
the European Union
NextGenerationEU



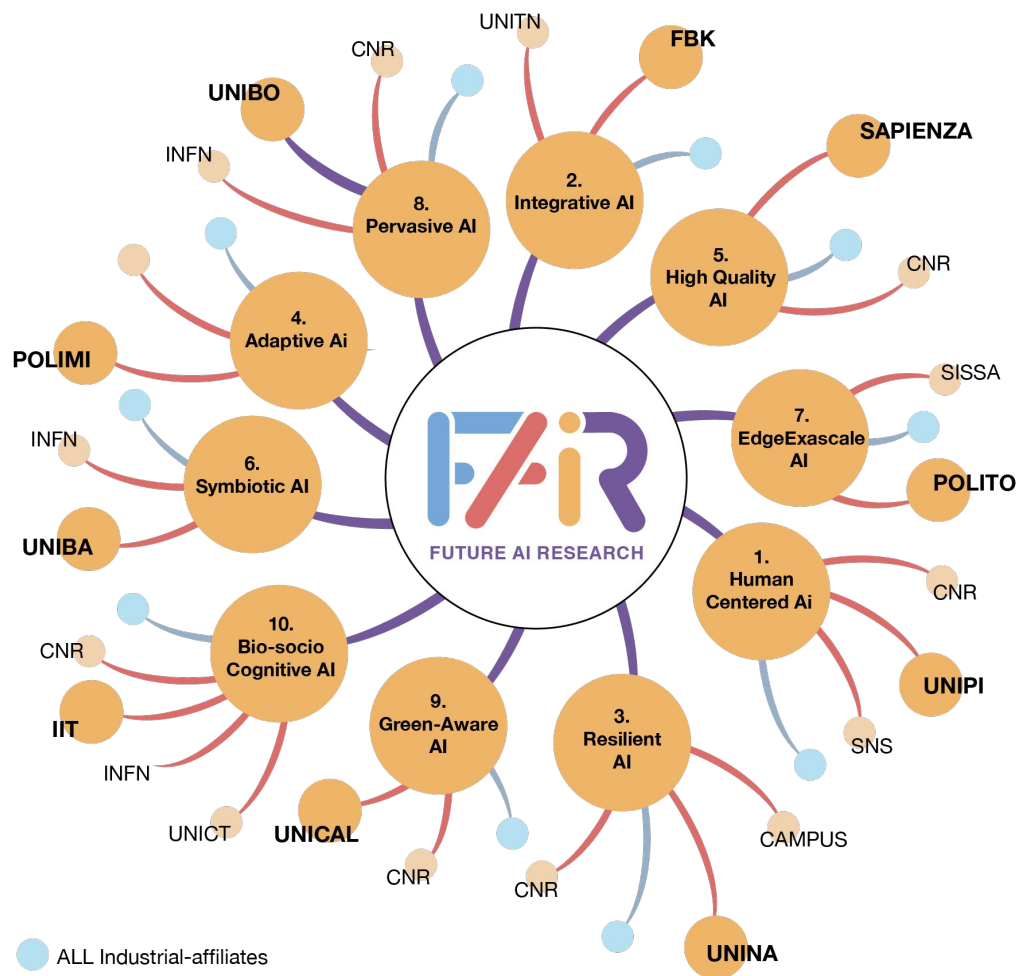
Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Future
Artificial
Intelligence
Research

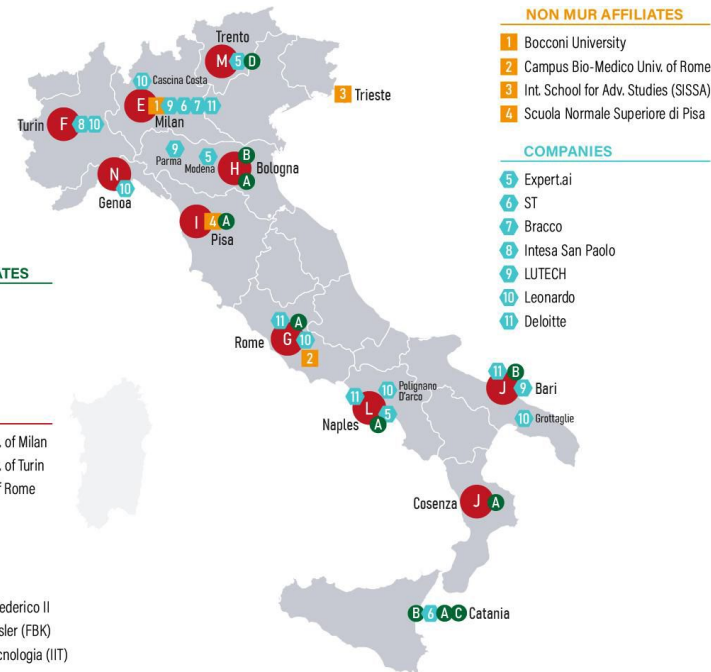


MUR AFFILIATES

- A CNR
- B INFN
- C Univ. of Catania
- D Univ. of Trento

SPOKE

- E Polytechnic Univ. of Milan
- F Polytechnic Univ. of Turin
- G Sapienza Univ. of Rome
- H Univ. of Bologna
- I Univ. of Pisa
- J Univ. of Bari
- K Univ. of Calabria
- L Univ. of Naples Federico II
- M Fond. Bruno Kessler (FBK)
- N Ist. Italiano di Tecnologia (IIT)



3 years: 2023-2025 Budget: 115 Meuro

Fig. A.2 Spokes and Affiliates



Future
Artificial
Intelligence
Research

Spoke 1 – Human-centered AI

Dino Pedreschi, Francesco Marcelloni
Università di Pisa

Spoke 1 - Human-centered AI - Partners



UNIVERSITÀ DI PISA



Consiglio Nazionale
delle Ricerche



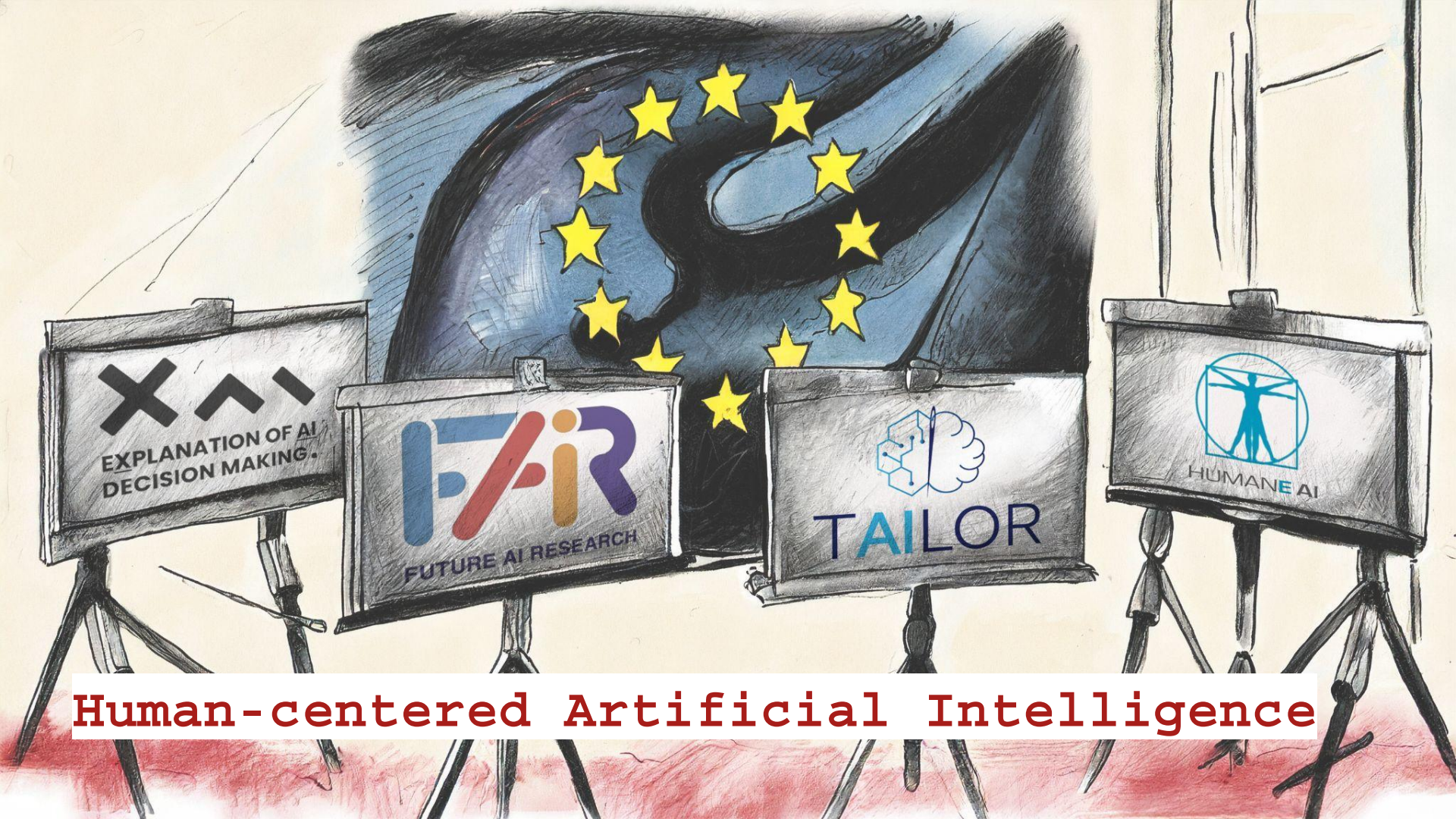
SCUOLA
NORMALE
SUPERIORE

INTESA



SANPAOLO

Deloitte.



Human-centered Artificial Intelligence

Human-centered Artificial Intelligence

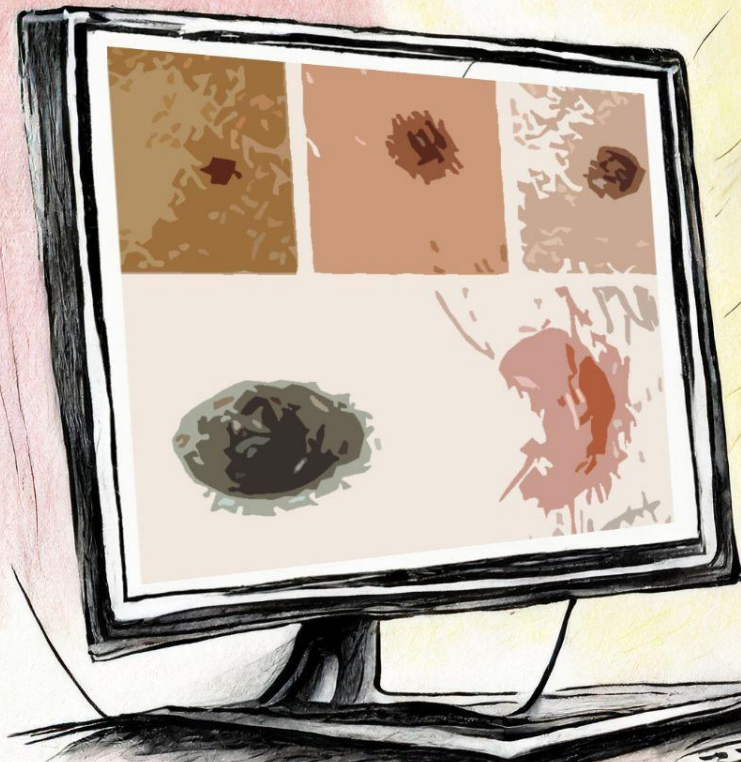
The study of **AI systems that cooperate synergistically, proactively and purposefully with humans** at individual and collective scale

- **amplifying instead of replacing human intelligence**
- maximizing benefits while preventing and minimizing risks.

Adaptive, collaborative, responsible, interactive, coevolving intelligent human-AI ecosystems.

Human-centered AI systems leverage human strengths and compensate for human weaknesses, while respecting **fundamental human rights** and taking into account the **social, ethical and legal context**.

enhancing rather than
replacing human
intelligence



→ Choose one of the case studies we selected using the menu below

id:156 - class:Melanoma



Image to explain (predicted class)

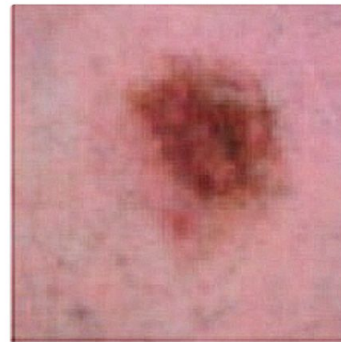


Melanoma

Neighborhood:

- Melanoma: 7
- Melanocytic nevus: 488
- Basal cell carcinoma: 181
- Actinic keratosis: 32
- Dermatofibroma: 194
- Vascular lesion: 98

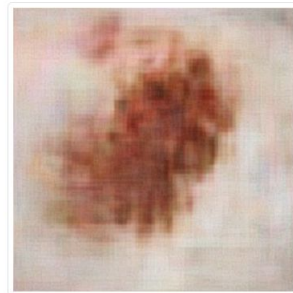
Counter example image (class)



Melanocytic nevus

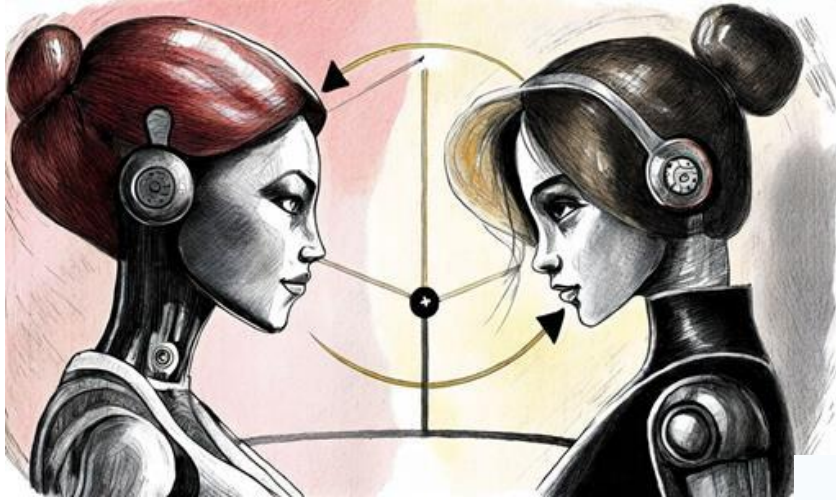
Prototype images

The following images are generated synthetically and they are classified with class **Melanoma** by the blackbox.

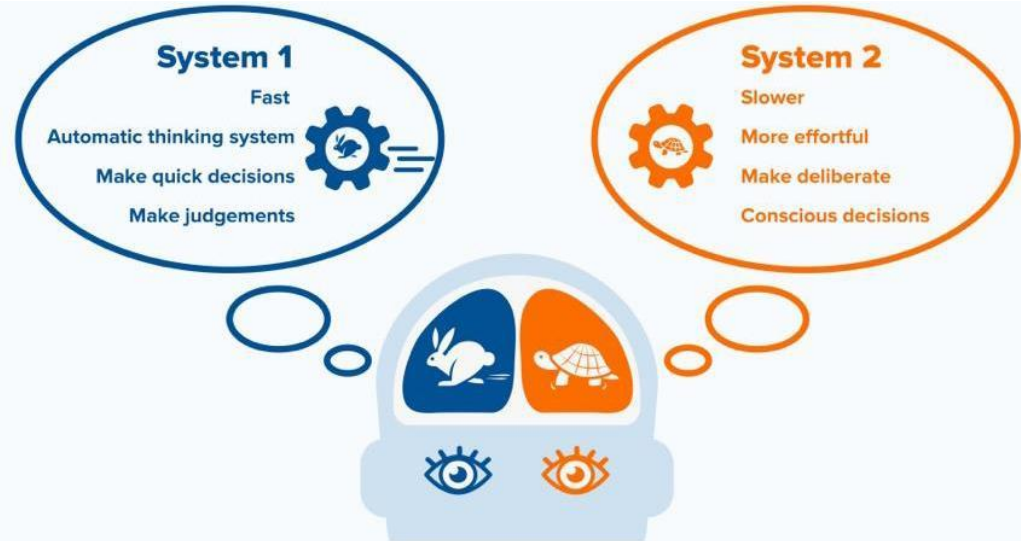




AI systems aware of their own limitations

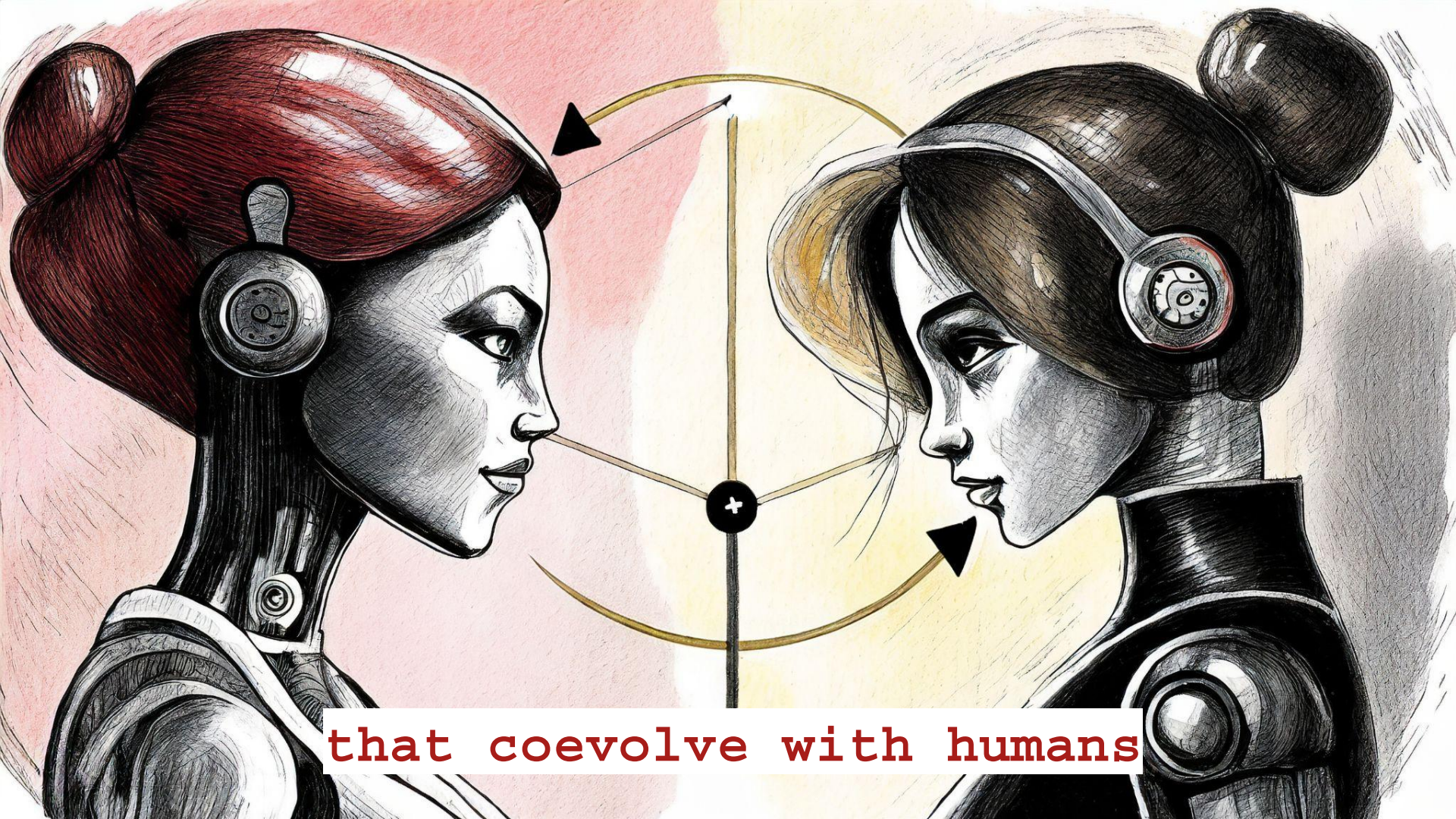


Decision making as a synergistic collaboration of human system1 and system2





thought partners



that coevolve with humans

Great game changers

Gen-AI and **LLMs**, that magnify both opportunities and risks (**human-like** doesn't mean **human-centered**)

EU's **AI Act**, the first systemic law on the AI market



Spoke 1: our research questions

1) **human-centered machine learning and reasoning:**

how humans and AI models interact synergistically, continuously co-evolving together (WP 1.1, 1.2, 1.3)

2) **human-AI ecosystems at societal scale:**

how to understand and govern the dynamics and societal outcomes of large-scale socio-technical systems of humans and AIs that coevolve together (WP1.4, 1.5)

3) **responsible design of trustworthy AI systems:**

how to responsibly (co-)design, develop, validate and use trustworthy AI systems (WP1.6)

Extensive **experiments, case studies and pilots** of H-AI systems (WP1.7).

WP1.1 - Explainable AI for synergistic Human-AI collaboration

Task lead: SNS, co-PIs: Fosca **Giannotti**, Riccardo **Guidotti**.

- **Task 1.1.1 Explanation of AI Black-Boxes**
- **Task 1.1.2 Design of Interpretable Models**
- **Task 1.1.3 Cognitive aware eXplainable AI for hybrid decision making**
- **Task 1.1.4 Decision Theory informed co-design of XAI Interfaces**
- **Task 1.1.5 En Plein Air Benchmarking & Human-in-the-loop**

The XAI-FAIR platform for eXplainable AI



FAIRNess
+ XAI

Abstaining
+ XAI

Frank

EnsembleC
FE

Visual XAI

TRACE

FIPER

CO-XAI Cognitive Decision Intelligence Framework for Explainable AI Systems

XAI and Defer impact on users: evaluation study

Joint project of SNS and UNIPi with Univ. Cattolica Milano

WP 1



Funded by
the European Union
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Future
Artificial
Intelligence
Research

WP1.2 – “System1 and System2” Machine Learning and Reasoning

Task lead: CNR, co-PIs: **Umberto Straccia**, **Salvatore Ruggieri**.

- **Task 1.2.1 Integrating neural and symbolic representations for learning and reasoning**
- **Task 1.2.2 Explainable neuro-symbolic systems**
- **Task 1.2.3 Exploiting Non-Standard Analysis to empower Deep Neural Networks and Reinforcement Learning**
- **Task 1.2.4 Mathematical structures in learning models**

WP1.3 - Human centered Lifelong Learning for Complex Data

Task Lead: UNIPi, co-PIs: **Alessio Micheli**, **Fabrizio Sebastiani**.

- **Task 1.3.1 – Foundations of human and pervasive AI**
- **Task 1.3.2 – Lifelong learning for socially and environmentally sustainable AI**
- **Task 1.3.3 – Learning in structured domains**
- **Task 1.3.4 – Cross-modal and cross-lingual learning**
- **Task 1.3.5 – Learning to quantify**
- **Task 1.3.6 – Improving Neural Language Models**
 - **see next highlight by Alessandro Lenci:**
 - **A native-Italian benchmark dataset to test the reasoning abilities of multimodal language models**

WP1.4 - Human-AI Socio-technical Complex Systems

Task lead: UNIPI, co-PIs: **Dino Pedreschi**, **Chiara Boldrini**.

- **Task 1.4.1: Models of society-scale, complex human-AI socio-technical systems**
- **Task 1.4.2: Social and economic impact of AI-STS (AI-assisted socio-technical systems)**
- **Task 1.4.3: Individual vs. collective goals of AI-STS**
- **Task 1.4.4: Self-organization and distributed information processing in AI-STS.**



Artificial Intelligence

Volume 339, February 2025, 104244

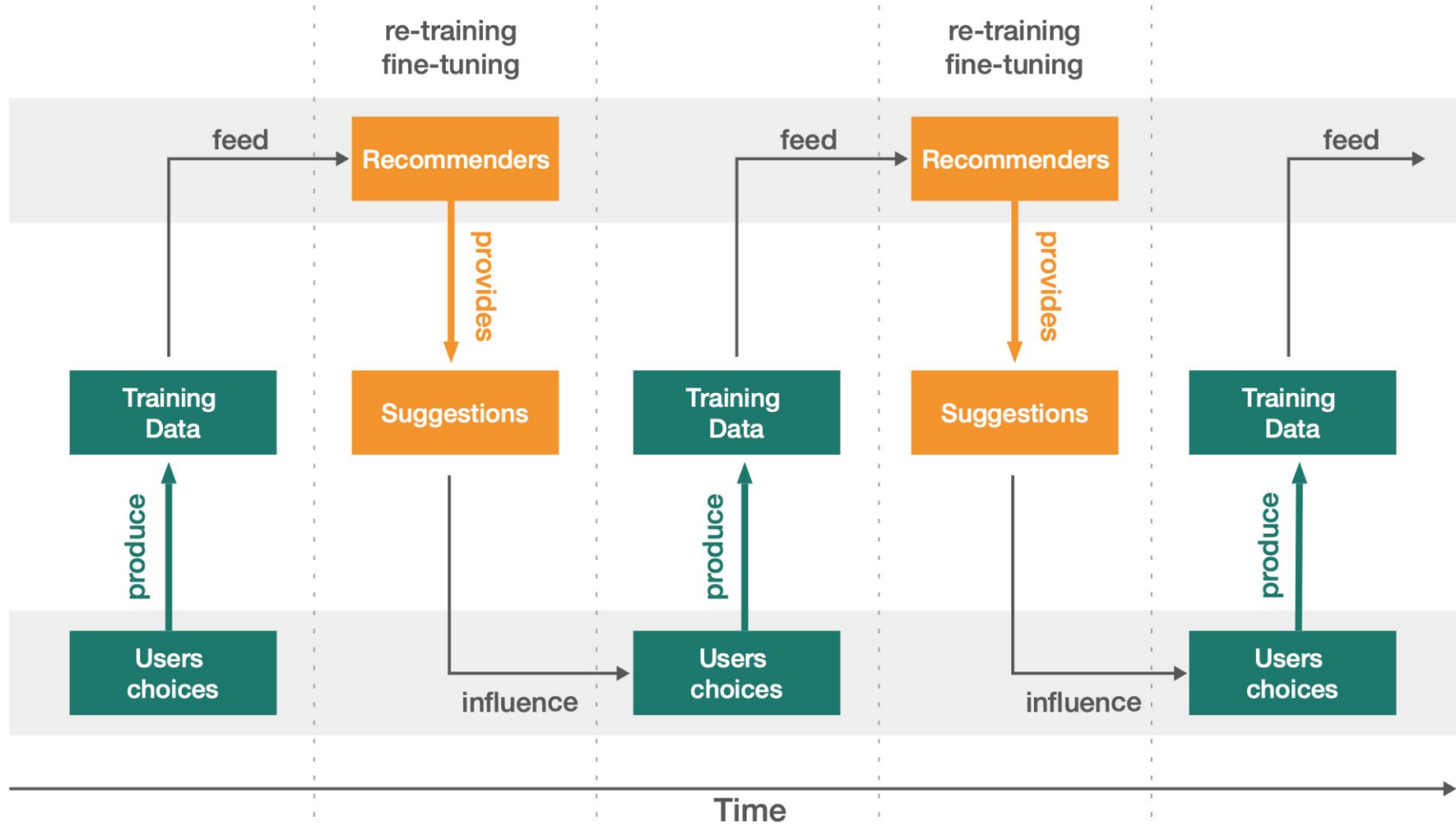


Human-AI coevolution

Dino Pedreschi ^a  , Luca Pappalardo ^{b c}  ,
Emanuele Ferragina ^d  , Ricardo Baeza-Yates ^e,
Albert-László Barabási ^e, Frank Dignum ^f, Virginia Dignum ^f,
Tina Eliassi-Rad ^e, Fosca Giannotti ^c, János Kertész ^g,
Alistair Knott ^h, Yannis Ioannidis ⁱ, Paul Lukowicz ^j,
Andrea Passarella ^b, Alex Sandy Pentland ^k, John Shawe-Taylor ^l,
Alessandro Vespignani ^e

A perpetual,
iterative process
wherein humans
and learning
algorithms
evolve in
tandem, each
influencing and
adapting to the
other over time.

THE HUMAN-AI FEEDBACK LOOP



WP1.5 – Decentralized, Cooperative Learning

Task lead: CNR, co-PIs: **Raffaele Perego**, **Francesco Marcelloni**.

- **Task 1.5.1: Novel fully distributed architectures for cooperative learning**
- **Task 1.5.2: Resource-efficient (non-centralised) learning**
- **Task 1.5.3: Robustness and control of non-centralised learning**
- **Task 1.5.4: Novel approaches to cooperative learning of XAI models**
 - **see next highlight by Francesco Marcelloni:**
 - **Federated Learning of Explainable AI Models: A Step Toward Trustworthy AI**

WP1.6 - Co-design methodologies for trustworthiness-by-design

Task lead: UNIPI, co-PIs: **Adriano Fabris**, **Anna Monreale**.

- **Task 1.6.1 Critical Oversight of EU Rules: limits and opportunities**
- **Task 1.6.2 Operationalization of Ethical and Legal Values**
- **Task 1.6.3 Ethics-for-designers methods for auditing and monitoring AI systems**
- **Task 1.6.4 Trustworthy-by-design AI systems**

Responsible AI & Generative AI

A course at the National PhD Program in AI

Anna Monreale, Salvatore Ruggieri (UNIFI),
Gizem Gezici, Fosca Giannotti (SNS)

WP 6



Funded by
the European Union
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Future
Artificial
Intelligence
Research

WP1.7 – Empirical studies and pilots of human centered AI

Task lead: UNIPi, co-PIs: **Francesco Marcelloni**, **Raffaele Perego**, **Salvatore Ruggieri**

3 conf. papers, 10 journals, 1 dataset

- **Task 1.7.1 :** Setup an open environment for large-scale empirical studies
- **Task 1.7.2:** Identification and definition of pilots and use cases
 - 7 Pilots
 - 14 Use cases
- **Task 1.7.3:** Pilots/Use Cases monitoring and development
 - 60+ researchers involved
- **Task 1.7.4:** Dissemination

Oral Squamous Cell Carcinoma (OSCC) Detection and Classification
Analysis of the effects of different diets on 24-hour core body temperature and energy metabolism
Engineering Relationship Recognition and Extraction
Extraction of State Machines from patent descriptions and images
Align of University Offers and Job Market with Natural Language Processing
Flash Simulation of High Energy Physics Events with Normalizing Flow techniques
The core and periphery of the scientific knowledge production in AI: a global network analysis.
Spatial Big Data: Exploring urban mobility through GeoAI.
Artificial Intelligence and new visions of the city: assemblages, imaginaries, policies
AI & socio-spatial urban inequality
Urban Archaeology: AI-controlled robotic arm for classification and selection of archaeological ceramic fragments
Urban Archaeology: Open AI solutions for historical masonry annotation (OPUS)
Environmental Archaeology: AI approaches for animal bones recognition and landscape reconstruction
Sub-IA: Project to interpret, digitally share, and document underwater archaeological heritage, as well as determine its ecological effects on local biological communities
Contemporary Archaeology: Digital traces of migration in Lampedusa
Industrial production quality recognition
Evaluation of an expressive social robot cognitive system by means of natural human-robot interaction and dialogue (ABEL)
Pancreatic ductal adenocarcinoma study with Explainable AI
Pre- and Post-Liver Transplant Diabetes (DM) and Prediabetes (PD), and the Impact on Transplant Outcomes: Results of a Large Single-Center Study.
Approaches to Debunk Fake News detected automatically
Quantifying query fairness under unawareness

UNIFI (25)
Bacciu Davide
Bernardeschi Cinzia
Cimino MGC Antonio
Cococcioni Marco
De Santis Federica
Ducange Pietro
Fabris Adriano
Fantoni Gualtiero
Gattiglia Gabriele
Greco Alberto
Guidotti Riccardo
Lazzerini Beatrice
Lazzeroni Michela
Lenci Alessandro
Lomonaco Vincenzo
Malizia Alessio
Marcelloni Francesco
Micheli Alessio
Monreale Anna
Neri Veronica
Pedreschi Dino
Pellecchia Enza
Rizzi Andrea
Romito Marco
Ruggieri Salvatore

SNS (5)
Benzi Michele
Giannotti Fosca
Pellungrini Roberto
Squartini Tiziano
Trenz Hans-Joerg

CNR (9)
Boldrini Chiara
Conti Marco
Fabrizio Falchi
Nanni Mirco
Passarella Andrea
Perego Raffaele
Rinzivillo Salvatore
Sebastiani Fabrizio
Straccia Umberto

Spoke 1 - Critical Mass
 39 multi-disciplinary scientists
 9 UNIFI Departments
 2 CNR Institutes
 2 SNS Classes

Young researchers hired by Spoke 1



Researchers

17



PhD students

11



Postdocs

13



Broadening the FAIR partnership in Spoke 1



**Cascade call for
basic research**

6 projects
running



**Innovation grant
for FAIR companies**

3 projects
running



**Cascade call for
companies**

11 projects
running

5 new academic partners



UNIVERSITÀ
CATTOLICA
del Sacro Cuore



UNIVERSITÀ
DI SIENA 1240

IMT

SCUOLA
ALTI STUDI
LUCCA



UNIVERSITÀ
DEGLI STUDI
FIRENZE



Sant'Anna
Scuola Universitaria Superiore Pisa

11 new industrial partners



Spoke 1 key numbers: publications and submitted manuscripts, March 2025

Journals

124



Conferences

168

Books

43



2nd FAIR Workshop on Human-Centered AI



Funded by
the European Union
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Future
Artificial
Intelligence
Research

A colourful, refreshing deep dive in cutting-edge human-centered AI research!

<http://whucaai2025.isti.cnr.it>

Two full days: 10-11 March 2025

Participants: 130

44 presentations (4 TPs, 2 Innovation grant,
6 academic partners, 7 use cases)

A blossoming ecosystem

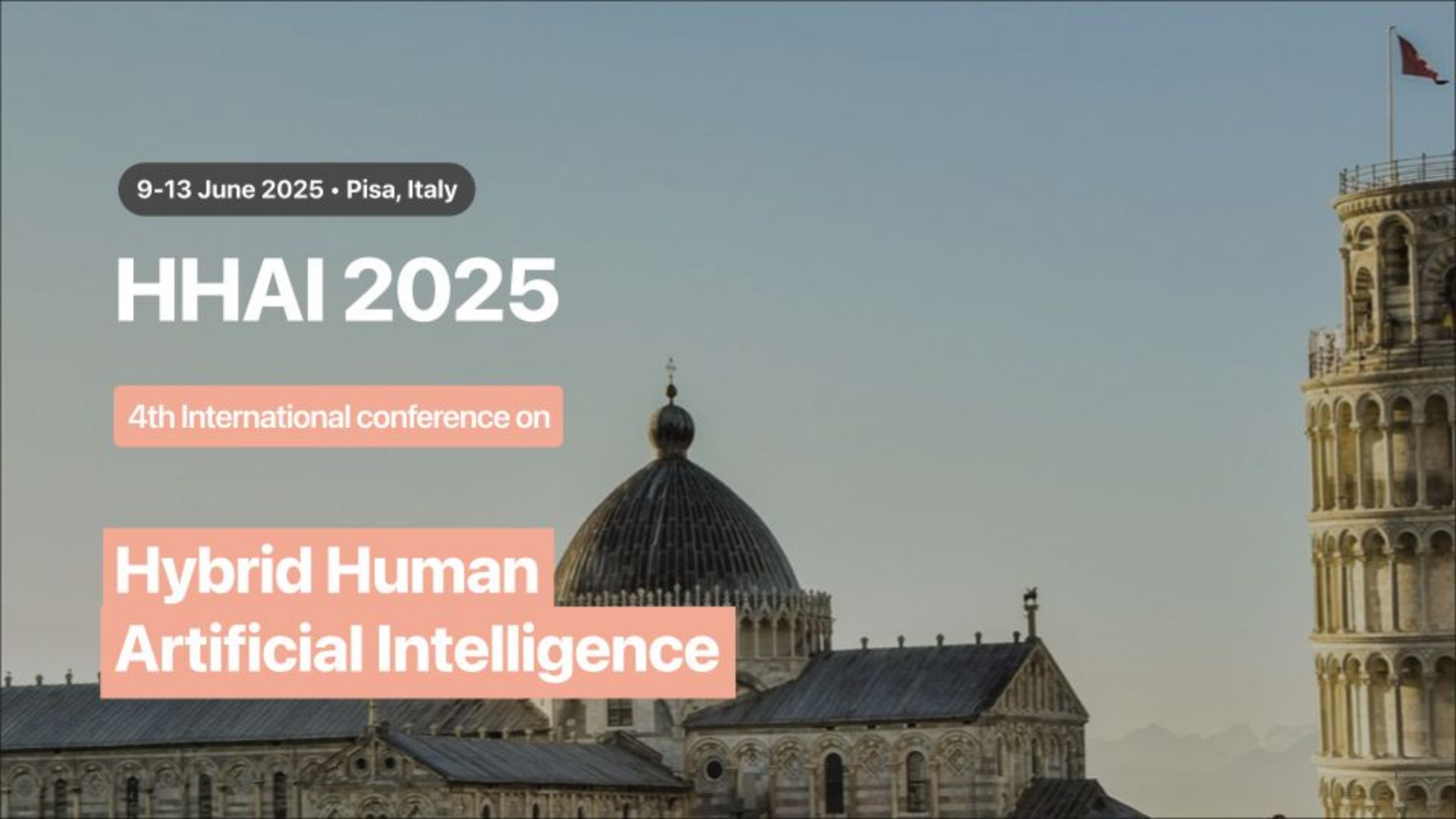
- Spoke 1 community is growing in size and complexity
- Impressive advance
- Both human-centered AI methodologies and use cases
- Multidisciplinary cross-fertilization

9-13 June 2025 • Pisa, Italy

HHAI 2025

4th International conference on

**Hybrid Human
Artificial Intelligence**



Conference Program

5 days of cutting-edge AI research

June 9-10

Pre-conference

14 Workshops

5 Tutorials

Doctoral Consortium

AI & Art Exhibition

June 11-13

Main Conference

4 Keynotes

7 Sessions

1 Interdisciplinary Panel

4 Poster & Demo Sessions

Session themes

Session 1



Human-AI Collaboration & Hybrid Intelligence

Session 2



Ethics, Trust & Moral Reasoning

Session 3



Human-AI Interfaces & Interaction

Session 4



Explainability, Perception & Decision Support

Session 5



Societal Impact & AI in Health, Crisis & Governance

Session 6



Adaptive Systems, Attention, and Diversity

Session 7



Generative, Creative & User-Centered AI

Jonathan Stray



UC Berkeley

How AI Might Make
Human Conflict
Better or Worse

June 11

John ShaweTaylor



University College
London & UNESCO
IRCAI

How AI Might Make
Human Conflict
Better or Worse

June 11

Stuart Russell



UC Berkeley

Human AI
Collaboration:
Perspectives and
Directions

June 12

Ricardo Baeza-Yates



Barcelona SC

The Limitations of
Data, Machine
Learning & Us

June 13

9-13 June 2025 • Pisa, Italy

<https://hhai-conference.org/2025/>

HHAI 2025

4th International conference on

**Hybrid Human
Artificial Intelligence**

