

EOSC SYMPOSIUM 2026 | 14-16 OCTOBER, FLORENCE

AI Readiness

Towards Rigorous and Trustworthy AI in Science

Mercè Crosas, Ph.D.

Director, Computational Social Sciences and Humanities, Barcelona Supercomputing Center
President, CODATA



With the support of
 **Generalitat de Catalunya**
Government of Catalonia



EuroHPC
Joint Undertaking



Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.



About our team: The CSSH Lab at BSC

The Barcelona Supercomputing Center (BSC)



EuroHPC
Joint Undertaking



Since 2005

The Computational Social Sciences and Humanities (CSSH) Interdisciplinary Initiative

Established in 2024, merging domain expertise with HPC, data, and AI engineering.

Studying Today's Society

Analyzing information ecosystems that shape choices in democracies:

- Public discourse & social media
- Public administration records
- Legal & regulatory frameworks
- Scientific knowledge
- Private data ecosystems

Studying the Past

Extracting historical insights using advanced computation:

- NLP for humanities
- Computational archaeology
- Computational social anthropology
- AI for cultural heritage & museums

Outline

01

AI in Science: The last decades

02

Science values and current challenges

03

**Preparing for AI in Science:
The European context**

04

**Advancing AI Readiness:
The AI4SOCIAL+ project**

05

Three Recommendations

Outline

01

AI in Science: The last decades

02

Science values and current challenges

03

Preparing for AI in Science:
The European context

04

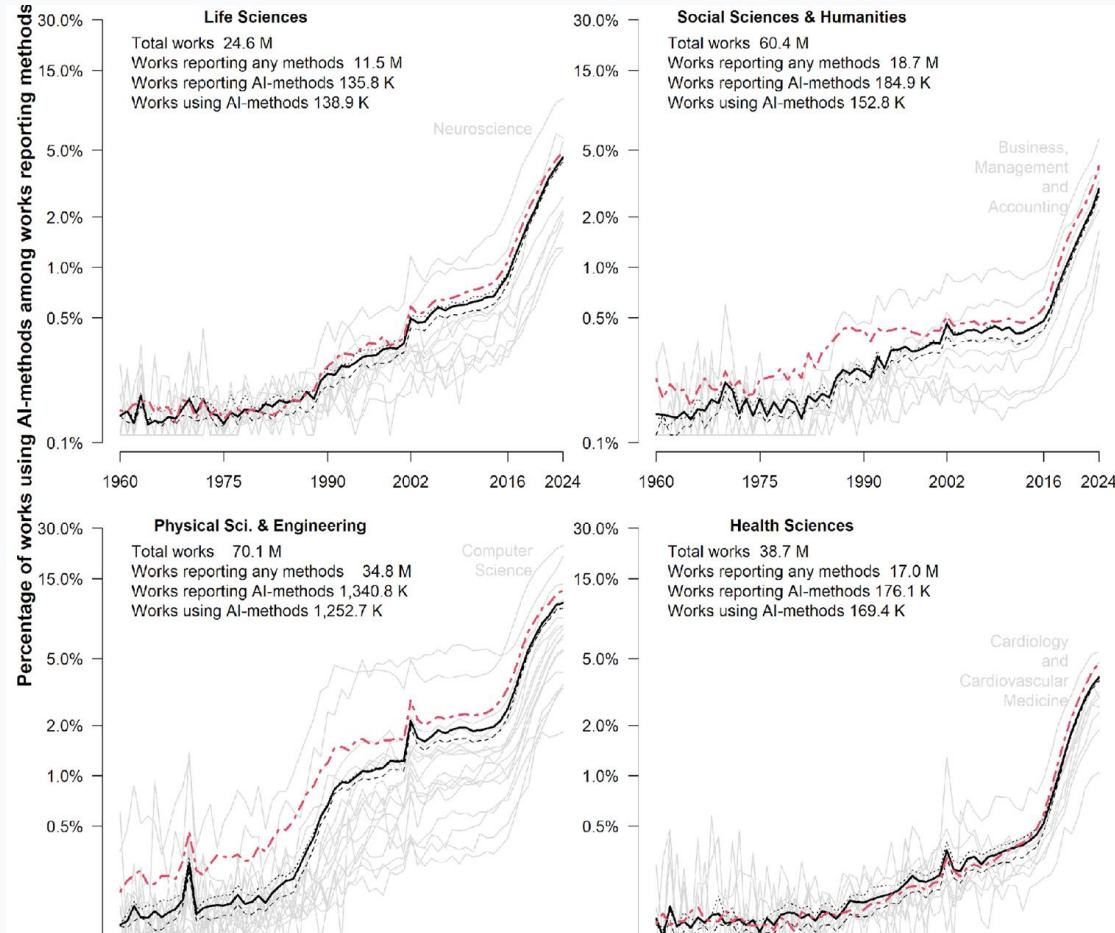
Advancing AI Readiness:
The AI4SOCIAL+ project

05

Three Recommendations

AI Adoption Surges Across All Domains

- **Study Scope:** Analysis of 225M academic works from 1960s to 2024
- **Long-term Trend:** CS leading since 1990s • **Acceleration:** Rapid expansion across fields since 2016



2024 AI Adoption Rate per Domain and leading Field

10.7%

Physical Sci. & Eng.

23.0%

Computer Science

4.4%

Life Sciences

10.0%

Neuroscience

3.7%

Health Sciences

5.4%

Cardiology

2.8%

Social Sciences

5.7%

Business Management

Spectrum of AI Engagement in Research

- **Broad AI Scope:** Defines AI broadly, beyond the latest Large Language Models (LLMs)
- **Study Priority:** Focus is explicitly on High-Engagement AI-adoption

Low-Engagement AI as a black box

- **Broad User Base**
- AI as a support tool
- Literature and data Searches
- Peer Reviews
- Research Reasoning

High-Engagement Better understanding of the AI model

PRIMARY FOCUS

- **Specialized Researchers**
- AI as method for analysis and discovery
- Foundation Model Development
- Model Modification (fine-tuning, RAG)
- Inference Execution

AI vs. Non-AI Research Patterns

Key findings from our study on AI-based research:

- **01. FOCUS:** Concentrated in fewer topics
- **02. METHODOLOGY:** Relies more on existing statistical frameworks
- **03. INTERDISCIPLINARITY:** Cites computer science more
- **04. IMPACT:** Receives more citations
- **05. INTEGRITY RISK:** Higher retraction rate

KEY IMPLICATION & CHALLENGE

Rapid AI growth requires safeguarding scientific integrity, open science, and community values.

Outline

01

AI in Science: The last decades

02

Science values and current challenges

03

Preparing for AI in Science:
The European context

04

Advancing AI Readiness:
The AI4SOCIAL+ project

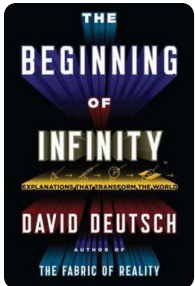
05

Three Recommendations

Scientific Theories & Social Practice

Scientific Theories are Good Explanations

Science is the pursuit of good explanations—marking the *beginning of infinity* in knowledge creation.



- **Good explanations arise from creativity**

Driven by creative insight, not solely induction and empiricism.

- **Fallibilism is fundamental**

Theories continuously evolve through conjecture, criticism, and testing.

- **Rejection of authority**

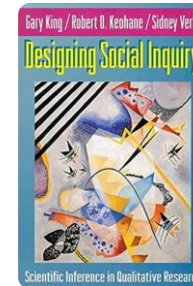
Nullius in verba: take no one's word for it.

- **Prediction is not enough**

True progress demands resilient explanations, not just data prediction.

Science is a Social Endeavor

Science is a **social enterprise** governed by norms and practices within scientific communities.



- **Public Procedures**

Methods and data must be open to enable verification and replication.

- **Accompanied Uncertainty**

Scientific conclusions are always presented with uncertainty estimates.

- **Unity in Method**

As Karl Pearson stated, unity in science consists in its method, not its subject.

- **Recognition through proper credit**

“Citation, Citation, Citation”.

Rapid AI-adoption poses threats to science and open science

Threats to Science

- **Disqualification**

Loss of core domain knowledge and subject understanding.

- **Disorder**

Surge of unverified and unreviewable research results.

- **Motivation Loss & Conformism**

Decline in creative human questioning and novel hypotheses.

- **Hyper-Empiricism**

Over-reliance on data without theoretical frameworks.

Threats to Open Science

- **Methodological Opacity**

Opaque workflows, models, and data, and lack of methodological transparency.

- **Replication Difficulty**

Obstacles in tracking provenance and reproducing workflows.

- **Open Data Exploitation**

Unaligned commercial scraping of public research data.

- **Big Tech Monopolization**

Public research infrastructure shifting into corporate silos.

Addressing threats starts by answering key questions

01. APPLICABILITY & PURPOSE

How and where should AI be applied in each scientific domain, and when does it truly outperform traditional methods?

02. VALIDATION & TRANSPARENCY

How do we validate AI results, assess uncertainty and bias, and ensure openness, documentation, and reproducibility?

03. DATA & GOVERNANCE

What conditions should we set for AI models and companies when utilizing research data and scientific repositories?

04. HUMAN AGENCY & EQUALITY

How do we keep scientists in the loop, safeguard critical thinking, and prevent power inequalities across regions?

Outline

01

AI in Science: The last decades

02

Science values and current challenges

03

**Preparing for AI in Science:
The European context**

04

Advancing AI Readiness:
The AI4SOCIAL+ project

05

Three Recommendations

AI & Data Regulations to Achieve AI & Data Strategies



Regulations addressing algorithm transparency, trustworthy AI, data protection, and sharing—appropriate, but risk lacking enforcement as AI expands.

DSA

Disclosure & (vetted) data access

Protects: Users, researchers, regulators

Status: Early Aug 2023, full Feb 2024

DMA

Behavioral obligations on gatekeepers

Protects: Business users, core services

Status: In force since May 2022

AI Act

Technical documentation along value chain

Protects: Deployers, end users

Status: Staggered from Aug 2026

GDPR

Personal data collection & processing rules

Protects: Personal data & privacy rights

Status: Enforcement started 2018

DGA

Neutral public data intermediaries framework

Protects: Data-sharing individuals & businesses

Status: Enforcement since 2022

Data Act

Data portability & fair access rules

Protects: Users from unfair contracts & data lock-in

Status: Partial 2025; Full 2026–2027

Parallel Initiatives advancing European AI Readiness

EuroHPC AI Factories

2024 Launch |

GigaFactories scheduled for 2027

Extends public supercomputing capacity to train, adapt trustworthy models for SMEs and public sector.



EuroHPC
Joint Undertaking



EOSC AI Integration

AI-related projects in 2026

Drives machine actionability and AI Readiness across the European Open Science Cloud ecosystem.



RAISE Initiative

2025 Launch

Brings core access to data, computing, talent, and funding. Funds specialized networks of AI scientific excellence.



Destination Earth

2022–2030 (DestinE)

Flagship Project using AI and digital twins to simulate Earth environmental and climate phenomena.



03. EUROPEAN CONTEXT

AI Factories and Data Labs

19 13

Core AI Factories Connected Antennas

- ✓ Compute (GPUs)
- ✓ Models, data services, user support, training, and expert advice based on specific domains.

Seven Sectoral Data Labs:

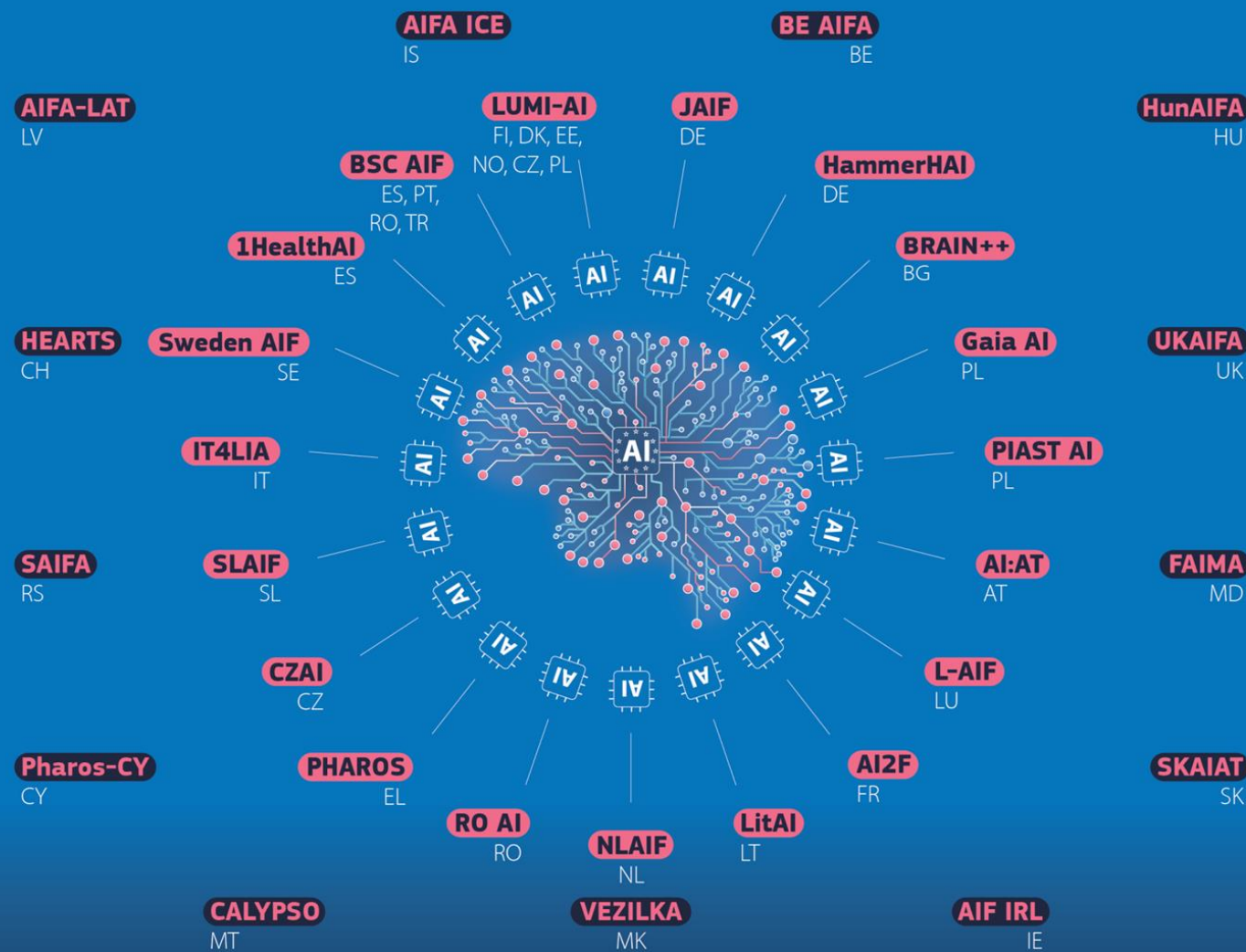
- **Cybersecurity, Internal Security**
- **Climate and Environment**
- **Language and Culture**
- **Health and Life Sciences**
- **Manufacturing and Robotics**
- **Public Administration**
- **Research and Science**

AI-ready data and connector from
Data Spaces to AI Factories



EuroHPC AI Factories (AIFs) and Antennas (AIFAs)

will enable European startups, SMEs, and researchers to develop, train and deploy AI solutions, and boost EU competitiveness and sustainable prosperity.



EuroHPC JU

LEADING THE WAY IN EUROPEAN SUPERCOMPUTING

Outline

01

AI in Science: The last decades

02

Science values and current challenges

03

Preparing for AI in Science:
The European context

04

**Advancing AI Readiness:
The AI4SOCIAL+ project**

05

Three Recommendations



AI Readiness for Social Impact

HORIZON-INFRA-2025-01 Research and Innovation Action

Coordinator: BSC (PI: Mercè Crosas) 14 Partners · 8 Countries · 36 Months | Budget: €7.49M (EU Funded)

Beneficiaries:



Affiliated Entities:



Why AI Readiness Matters for the Scientific Community

AI Readiness in Data & Research Workflows

- Enables data stewardship & best practices for AI data/code.
- Emphasizes documenting (FAIR) data, models, and workflows.
- Guides development of infrastructure and AI research tools.

AI Readiness in Science Norms & Values

- Addresses decisions on AI made by research communities.
- Motivates the community to redefine values that guide their research.
- Integrates ethical standards into AI adoption.

AI4SOCIAL+ Core Aims

AIM 01

1. Advance AI Readiness & Actionability

Advance AI readiness and machine actionability within the **EOSC ecosystem** and broader research networks, ensuring a rigorous, trustworthy, and standardized approach.

In collaboration with:

 FLUID-AI |  EOSC

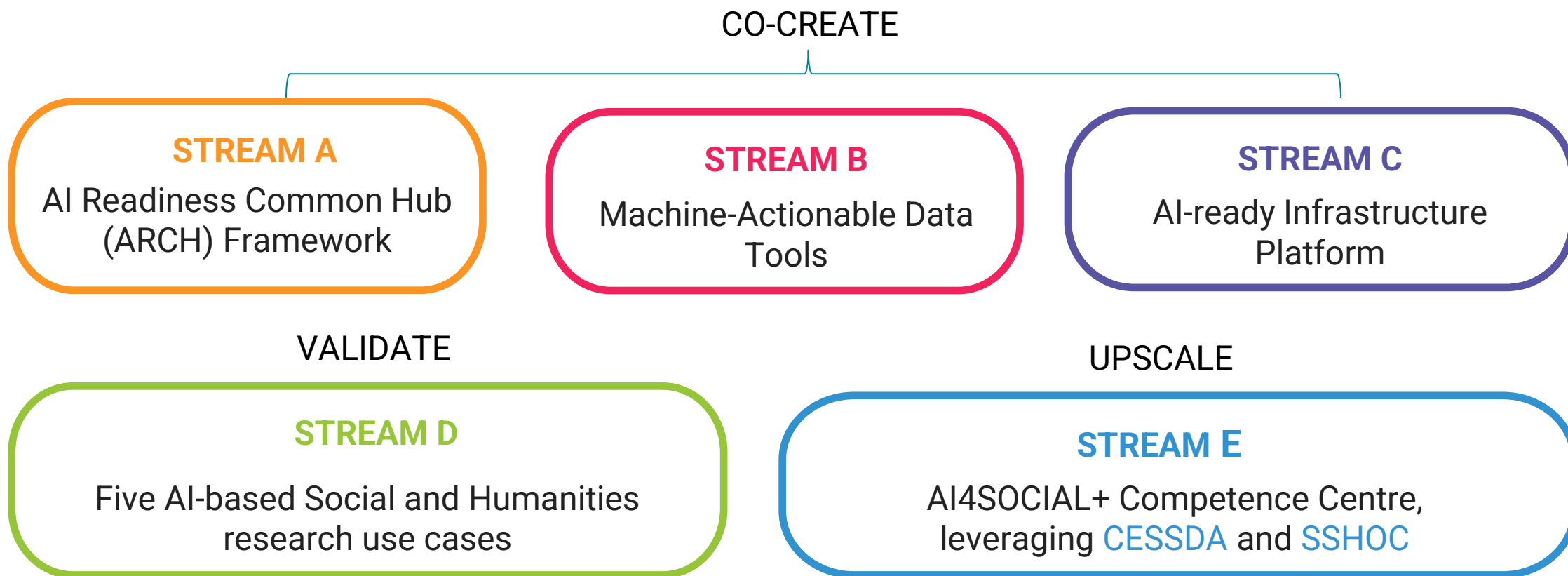
 CDIF 4 |  EOSC

AIM 02

2. Equip the Social Sciences & Humanities

Equip and elevate the SSH domain to adopt AI in strategic areas offering the highest potential for domain transformation and direct social impact.

Five Interrelated Streams to Advance AI Readiness



Building upon existing standards, tools, and platforms:



Croissant



RO-Crate



Five research Use Cases to Validate AI Readiness

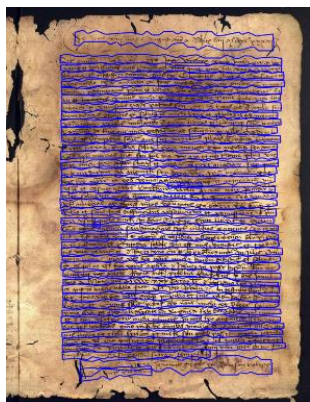
The AI-ready Framework, Tools, and Infrastructure tested against research with wide social impact



Use Case 01

AI for Job Trajectories

Analyzing career paths, labor market dynamics, and skill transitions at the individual level.



Use Case 02

AI in Historical Archives

Transcribing and unlocking insights from extensive cultural and historical records.



Use Case 03

AI for Climate & Health Policies

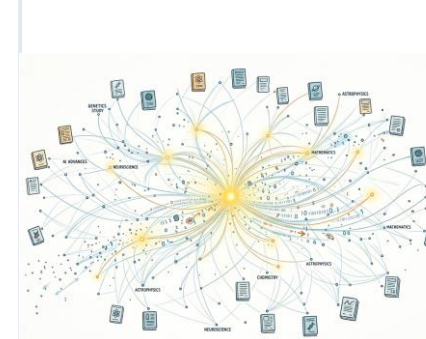
Informing policy choices at the intersection of public health and environment.



Use Case 04

AI for Social Simulations

Modeling complex social behaviors and interventions with multi-agent models.



Use Case 05

Beyond Academic Silos

Enabling cross-domain collaboration by connecting works through AI Graphs

Outline

01

AI in Science: The last decades

02

Science values and current challenges

03

Preparing for AI in Science:
The European context

04

Advancing AI Readiness:
The AI4SOCIAL+ project

05

Three Recommendations

Recommendation 1: Coordinate AI efforts Across EU Initiatives

EU has the necessary Ingredients:

Data and AI Regulations & Initiatives

The ingredients exist in Europe to create a rigorous, trustworthy, and sustainable global ecosystem that benefits science and society.

Key Initiatives to Align:

- EOSC
- RAISE
- AI Factories & Data Labs
- Destination Earth

Recommendations to EOSC for alignment:

- **Coordinated Initiatives and Calls:** EOSC, RAISE, and AI Factories together on selected key issues.
- **Connected Data Repositories and Data Access:** EOSC data infrastructures, Data Labs, Data Spaces
- **Standardized AI Readiness across EU initiatives:** Leverage AI4SOCIAL+ AI Readiness Framework.
- **Sustainable AI:** Integrate with Destination Earth for green AI models and AI limits.
- **Co-located Events:** Promote shared events for cross-initiative collaboration.
- **Targeted Research Funding:** Study AI impacts on science, equality, labor, and social interactions.

Recommendation 2: Co-define Declarations on AI and Science

The Precedent: Leiden Declaration on AI & Mathematics (2026)

"We base recommendations on characteristic values of mathematical research we aim to preserve."

Disclose automated tools, retain human accountability, defend open science, and preserve authorship credit.

<https://leidendeclaration.ai>

Plan from AI4SOCIAL+ consortium:

- Co-create Declarations on AI and Social Sciences and Humanities.

Call for action to scientific communities:

- Avoid becoming passive AI users – active voice.
- Decide how to integrate AI into research.
- Express what type of AI your domain wants & needs.
- Tackle community-specific needs.
- Find commonalities with other scientific domains
- Ensure long-term disciplinary flourishing.
- Address AI challenges while embracing benefits.
- Consider science impacts and global contexts.

Recommendation 3: Enforceable Data Contracts and APIs for AI

The Precedent: Wikimedia Enterprise Model

Wikimedia halted open scraping by AI tech commercial organizations, establishing paid API services via contracts for high-volume usage, while ensuring public content remains completely free for other users.

Potential access conditions for AI Models:

- Restrict to scientific AI models.
- Require trustworthy AI compliance.
- Allow only sustainable (green) AI.

Invite Research Data Repositories to define Data Contracts for AI and enforceable mechanisms:

- Balance Open Science with commercial AI usage limits and contracts.
- Create dedicated AI-purpose APIs while preserving open researcher access.
- Embed technical enforcement directly within repositories.
- Advocate for supportive regulatory frameworks.
- Enforce strict model-level access criteria.
- Leverage Wikimedia open-source API services.

Thank you!

- **Balanced AI Integration:** AI is transforming scientific research rapidly, requiring a balance between maximizing opportunities and mitigating risks.
- **Standardized AI Readiness:** The AI4SOCIAL+ project helps define and standardize AI readiness framework to deliver lasting value for the scientific community and society.
- **Coordinated EU Initiatives:** Regulations alone are insufficient; strategic alignment across computing, data, skills, and funding across the EU is vital.
- **Community AI Declarations:** Research communities must establish declarations on AI to preserve disciplinary values, credit, and human accountability.
- **Enforceable Data Contracts:** Repositories should implement data contracts and structured APIs to regulate AI training access on research data, balancing with open science.