ACHILLES

HUMAN-CENTRED MACHINE LEARNING:
LIGHTER, CLEARER, SAFER

Spotlight on Early Achievements: First insights into AI fairness metrics and stakeholder engagement

ACHILLES unveils first results in AI fairness metrics and stakeholder-driven development

The ACHILLES project — funded under the Horizon Europe programme and launched in November 2024 — is proud to share its **first early results** in the field of **trustworthy and compliant Artificial Intelligence (AI)**, marking a significant step forward in the pursuit of "**Lighter, Clearer, Safer" machine learning.**

Focusing on **AI fairness, data quality, and stakeholder engagement**, the project has made important progress in the definition and operationalisation of **fairness metrics**, the development of bias detection and mitigation tools, and the validation of **user-centered design methodologies.**

Key Early Results:
Metrics for Fair AI and Human-Centric Design

Fairness Metrics & Bias Toolkits:

ACHILLES has released preliminary frameworks and toolkits to **identify, assess and mitigate bias** in machine learning datasets and models. These include **group fairness and counterfactual fairness metrics**, developed under WP1 and WP5, with validation across real-world use cases such as healthcare diagnostics and identity verification.

Data Auditing & Minimum Viable Datasets:

Partners have implemented novel techniques to create **Minimum Viable Datasets (MVDs)**, enhancing data quality while promoting inclusiveness and representativity, critical for legal and ethical compliance.

Stakeholder Engagement & Value-Based Co-Design:

Early stakeholder workshops and user studies have already demonstrated the importance of **value-sensitive design and ethical AI impact assessments**. These activities are helping to align AI development with the values and needs of target communities, ensuring transparency and accountability from the ground up.



Consortium Perspective

> These initial results validate our vision: that AI systems can — and must — be designed with fairness, inclusivity, and accountability at their core.

Dr. André V. Carreiro, *Coordinator, Fraunhofer Portugal AICOS*

> Stakeholder input is not just a phase — it's an ongoing driver of responsible innovation. With ACHILLES, we are redefining co-creation in AI.

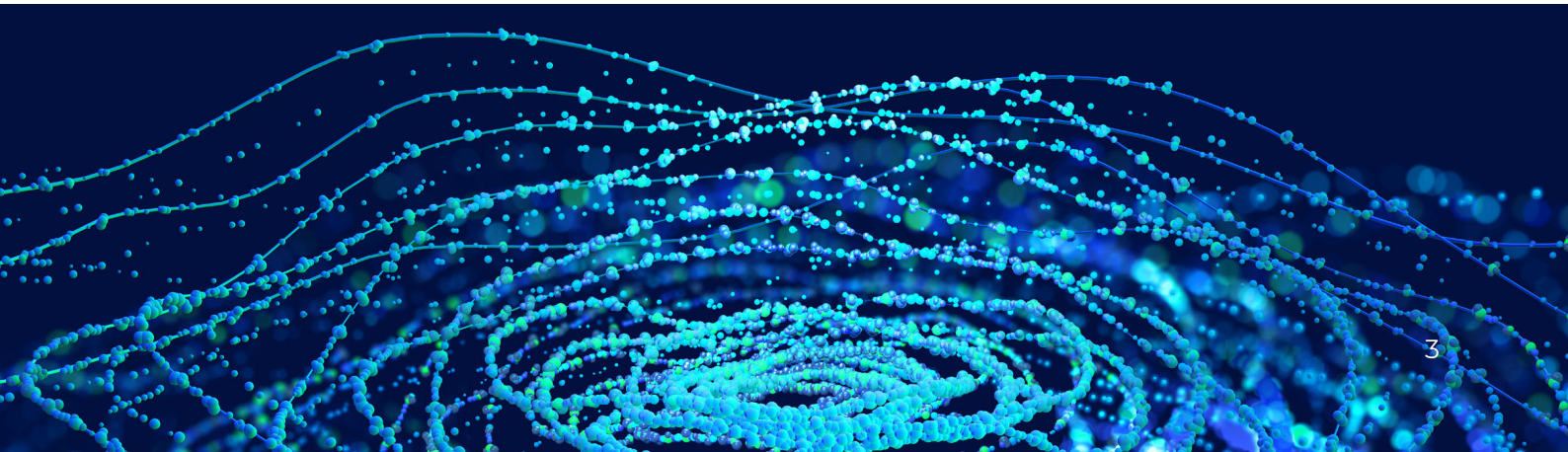
Silvia Colella, *Dissemination Lead, PNO*

Next Steps

In the months ahead, ACHILLES will:

- Finalize and publish Version **2 of the Auditing & Validation Toolkits**
- Expand its **use case testing** in medical, content creation, and pharma sectors
- Launch new features in the **ACHILLES IDE**, including privacy-preserving data handling and model explainability modules.

The consortium also invites external stakeholders to participate in upcoming workshops and public events.



Consortium

ARCADA

AXIOLOGIC

CTT-INNOVATIONS

Fraunhofer PORTUGAL

Fraunhofer HHI

IDnow.

inesc id lisboa

KU LEUVEN CiTiP
CENTRE FOR IT & IP LAW

PNO INNOVATION

INCM

Instituto de Sistemas e Robótica | LISBOA

IdISSC INSTITUTO de INVESTIGACIÓN SANITARIA Hospital Clínico San Carlos
Hospital Clínico San Carlos SaludMadrid

UNIVERSIDADE DA CORUÑA

eticas

CUOMO IT Consulting GmbH

www.achilles-project.eu

FOLLOW US

www.linkedin.com/company/achilles-project

ACHILLES
HUMAN-CENTRED MACHINE LEARNING: LIGHTER, CLEARER, SAFER

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.