



UNIVERSITÀ DI PISA



Advancements in Trustworthy Artificial Intelligence



Modern Artificial Intelligence (AI) models, such as Deep Learning (DL) and Neural Networks (NNs), are characterized by very high accuracy in solving predictive tasks but they often fail to address the growing need for **trustworthiness**.

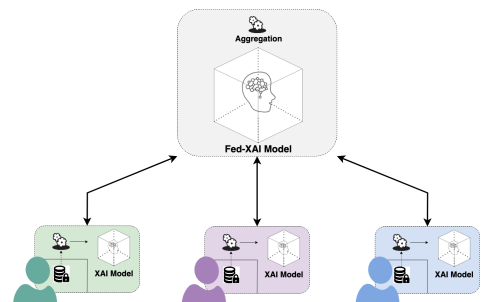
The requirement of trustworthiness in AI comprises several aspects:

- Protection of privacy of data owners in the learning stage
- Explainability, i.e. capability of providing explanations about decision
- Technical robustness, especially when data come in a streaming fashion

The candidate is expected to design and develop algorithms and techniques for meeting such requirements and to investigate their applicability in the framework of Trustworthy AI. The candidate will join a team with great experience in this topic.

Have a look at the following online sources:

- Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong. 2019. **Federated Machine Learning: Concept and Applications**. *ACM Trans. Intell. Syst. Technol.* 10, 2, Article 12 (March 2019), 19 pages. <https://doi.org/10.1145/3298981>
- Bárcena, José Luis Corcuera, et al. "Fed-XAI: Federated Learning of Explainable Artificial Intelligence Models." *3rd Italian Workshop on Explainable Artificial Intelligence (XAI. it 2022)*. 2022. <https://ceur-ws.org/Vol-3277/paper8.pdf>
- Online repository: <https://github.com/Unipisa/OpenFL-XAI>



Contact points:

- francesco.marcelloni@unipi.it
- pietro.ducange@unipi.it

Check the [AI-DII Group](#) website for other proposals

