



FAIR-PreSONS NEWSLETTER



FAIR-PreSONS: Advancing Fair AI in Recidivism Prediction

Fair predictions of gender-sensitive recidivism!

Welcome to the FAIR-PreSONS NEWSLETTER!

In this first issue, we share a brief overview of the project, its challenges, and the next steps relating to this exciting European project.



Brief Overview

FAIR-PreSONS, a DG JUST project, funded by the European Union and coordinated by the University of the Aegean (GR), aims to address and mitigate bias in recidivism data to ensure fair and equitable decisions in the criminal justice systems at the national and European level from a gender perspective. The project focuses on the collection and digitization of data from prisons and offense management systems (OMS) across participating countries, Greece, Portugal and Bulgaria, structuring this data into knowledge graphs, and making it accessible via a dedicated data portal. A preliminary gender analysis will be conducted to identify and address gender biases in recidivism prediction systems. The project employs advanced algorithms and integrates knowledge graphs with artificial neural networks (ANNs), aiming to enhance the accuracy and explainability of predictions. This holistic approach to fair AI aspires to support unbiased decision-making, thereby assisting judges and legal practitioners in their effort to make unbiased/fair decisions and mitigate gender stereotypes.



ATHENS
LIFELONG
LEARNING
INSTITUTE



CENTER FOR
THE STUDY OF
DEMOCRACY



INNOVATIVE
PRISON SYSTEMS



Co-funded by the
European Union

This project [101160473 – FAIR-PreSONS – JUST-2023-JACC-EJUSTICE] has been funded with support from the European Commission. This publication (communication) reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



Challenges

The main challenges of the project are:

- Collection and anonymization of appropriate data from penitentiary institutions and public bodies
- Research in AI and bias-free predictions
- Recidivism analytics
- (AI) Ethics
- Involvement of judges and legal practitioners

Next Steps

The initial phase involves the collection of data from national public agencies, ensuring a shared understanding and definition of recidivism. A technical meeting is scheduled to clarify concepts and streamline data requests. Concurrently, the preliminary gender analysis will commence to identify gender-based disparities and integrate gender-sensitive variables into the recidivism prediction system. The project will also investigate state-of-the-art algorithms and integrate knowledge graphs with ANNs to enhance prediction accuracy and support algorithm explainability. Efforts will be directed towards mitigating biases in data and developing tools for interpretability.

The engagement and training of end-users, including legal professionals and judges, will be a key focus. This involves identifying and engaging 80 participants with an emphasis on gender balance, and designing training sessions that address both ethical considerations and practical applications of the AI tools. A rigorous evaluation framework will be established to assess the AI tool's accuracy, fairness, transparency, and privacy safeguards. Ensuring continuous feedback and adherence to ethical standards will be integral to this process. Through these measures, FAIR-PreSONS is committed to developing a fair, unbiased, and transparent recidivism prediction system, thereby setting a precedent for ethical AI in the criminal justice system.

On behalf of the consortium,

The Coordinator



For further details please visit our Website
<https://fair-presons.aegean.gr/>



ATHENS
LIFELONG
LEARNING
INSTITUTE



CENTER FOR
THE STUDY OF
DEMOCRACY



INNOVATIVE
PRISON SYSTEMS



Co-funded by the
European Union

This project [101160473 – FAIR-PreSONS – JUST-2023-JACC-EJUSTICE] has been funded with support from the European Commission. This publication (communication) reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.