

AI4Gov - Zaupanja vredna umetna inteligenca za transparentno javno upravljanje, ki spodbuja demokratične vrednote

Alenka Guček

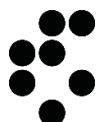
alenka.gucek@ijs.si

Kako nasloviti »Enakost spolov in vključevanje dimenzije spola« v projektni prijavi programa Obzorja
Evropa?

29.3.2023



Funded by
the European Union

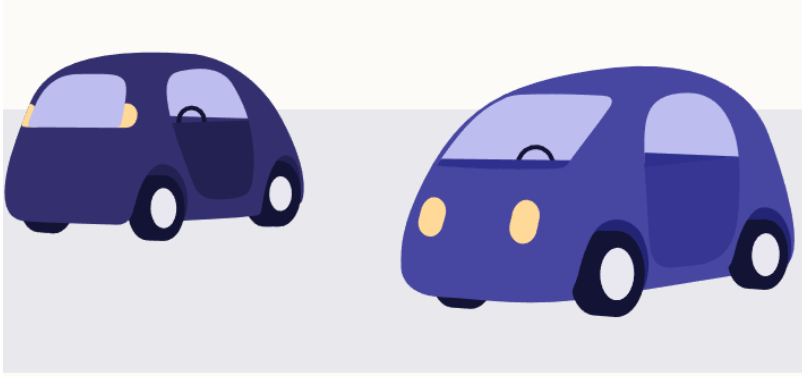


Institut
Jožef Stefan

Odsek za
Umetno Inteligenco



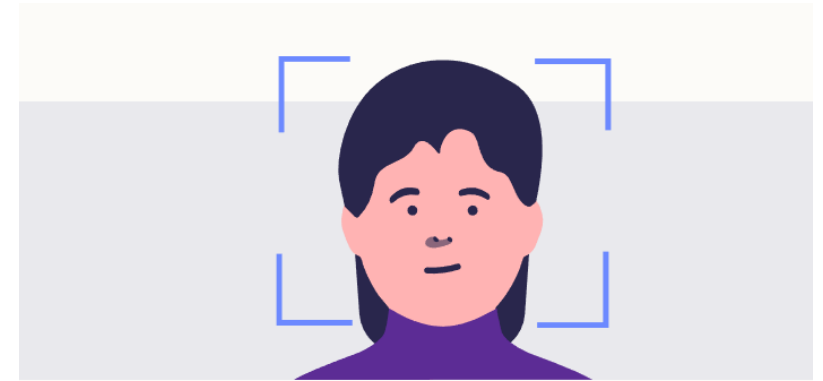
“Umetna inteligenca je simulacija procesov človeške inteligence z napravami.”



Uporaba 1: avtomobili brez voznika

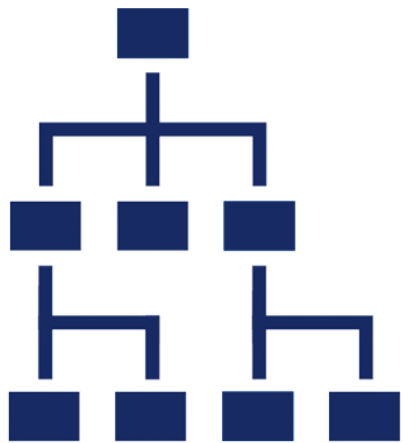


Uporaba 2: priporočanje vsebin



Uporaba 3: obdelava slik in videoposnetkov

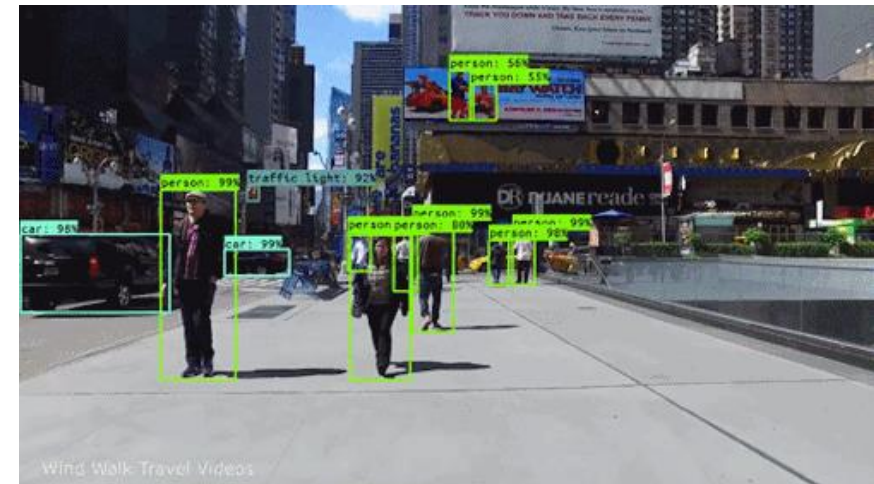
● Elements of AI



ekspertni sistemi

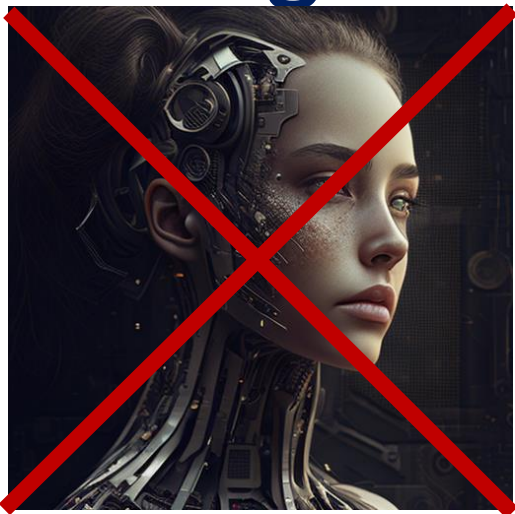


procesiranje naravnega jezika



računalniški vid

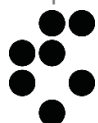
Algoritmi in pristranskost



Midjourney generirana slika, prompt: artificial intelligence



AI Research SuperCluster (RSC), Meta



Institut
Jožef Stefan

A screenshot of a news article from The Guardian. The article title is "Amazon ditched AI recruiting tool that favored men for technical jobs". A yellow banner above the title states "This article is more than 4 years old". The sub-headline reads: "Specialists had been building computer programs since 2014 to review résumés in an effort to automate the search process". Below the text is a photograph of a group of people standing in front of a building entrance with a sign that says "recruiting". A large yellow sign in the foreground features the Amazon logo and the text "FTJobsNow.com".

The Guardian

inion Sport Culture Lifestyle More ▾

Climate crisis Environment Science Global development Football Tech Business Obituaries

This article is more than 4 years old

Amazon ditched AI recruiting tool that favored men for technical jobs

Specialists had been building computer programs since 2014 to review résumés in an effort to automate the search process

recruiting

FTJobsNow.com

amazon

Amazon's automated hiring tool was found to be inadequate after penalizing the résumés of female candidates. Photograph: Brian Snyder/Reuters

Odsek za
Umetno Inteligenco



Trusted AI for
Transparent Public
Governance
fostering
Democratic Values



Holistic Regulatory Framework (AUTH)

A comprehensive analysis of multiple types of bias based on different grounds such as: age, disability, gender, race, sexual orientation, and gender identity will establish the grounds for the realisation of the project’s outcomes providing support for “regulatory compliance by design”.

Bias detector toolkit (IJS)

This toolkit will facilitate the extraction, simulation, evaluation, and optimization of enhanced evidence-based policies, for transparent, unbiased, and discrimination-free AI and Big Data applications.

Explainable AI (XAI) Library (IBM)

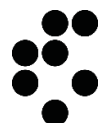
The XAI library will be used not only for explaining and understanding the underlying AI mechanisms, but also, for fostering the bias detection within the project.

1 (CO)		Maggioli S.p.A. [MAG]	Italy	Large Industries
2		SIEMENS S.R.L. [SIE]	Romania	
3		IBM Israel Science & Technology Ltd [IBM]	Israel	
4		Ubitech Limited [UBI]	Cyprus	SMEs
5		VILABS Limited [VIL]	Cyprus	
6		University of Piraeus Research Centre [UPRC]	Greece	RTOs
7		Institut “Jožef Stefan” [JSI]	Slovenia	
8		Aristotle University of Thessaloniki [AUTH]	Greece	
9		Comune Di Genova [CDG]	Italy	End Users - Policy Makers
10		Municipality of Vari, Voula, Vouliagmeni [VVV]	Greece	
11		Greek Ministry of Tourism [MT]	Greece	
12		White Label Consultancy ApS [WLC]	Denmark	Legal

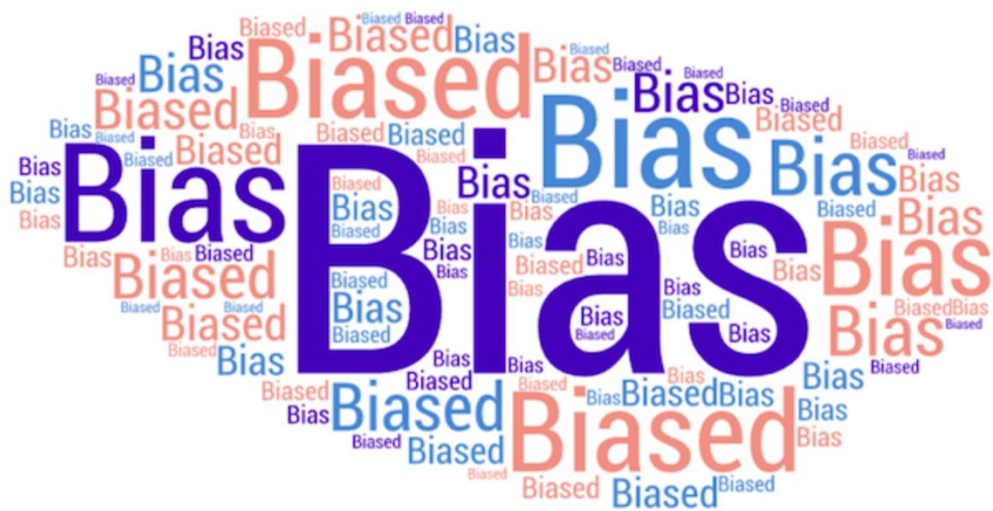
Načrt za enakost spolov

AI4Gov will adopt a **methodology that aims at conducting intersectional and interdisciplinary gender analysis** that will run through and beyond the project's scope, examining and understanding the effect of gender in AI, Big Data and democracy. An effective gender strategy will be used based on VILABS experience, tailored to the needs of the project, and thus increase its societal relevance and acceptance¹⁹ and fostering the gendered innovation process²⁰. Our methodological approach reflects on the 2 basic questions below for mainstreaming the gender dimension in the research process²¹: “*how gender differences and inequalities will affect project results and outcomes*” and “*how the project will differentially affect men's and women's health, opportunities, and potential intervention status*”. AI4Gov will collect primary and secondary gender-focused data from official and non-official sources and databases related to our objectives. The gender analysis will use a combination of quantitative and qualitative methods to fully assess the comparative and relational aspect of gender. The gender aspect as a cross cutting dimension will be tackled in WP 1 and will be systematically considered across all phases of the project, including:

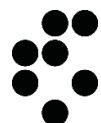
- **1) Tools for data integration:** AI4Gov will deliver and incorporate to the HRF various tools, such as XAI Library and Bias Detector Toolkit, that will prevent gender bias and discrimination within the project (see subsection 1.1.2).
- **2) The UCs Implementation Phase:** During the UCs implementation, the analysis of AI results/outcomes by data scientists and investing practitioners will consider their gender-specific aspects.
- **3) The Evaluation Phase:** Methods of the design process will socio-scientifically be complemented by gender-specific aspects. This is important as there may be gender-related differences in AI-based recommendations in terms of risk tolerance investments and commercial needs.
- **4) AI/ML models' Development:** Gender is one of the most prominent features towards unbiased and accurate AI/ML models. To this end, feature engineering will be applied during AI/ML models' development based on several characteristics (i.e., nationality, age, etc.), including gender. Gender factors will therefore consider in feature engineering processes for AI/ML development.



Intersekcionalnost in druge pristranskosti



Kar se ne šteje, ne šteje



Aktivnosti IJS na področju enakosti spolov

- Institut "Jožef Stefan" (IJS) v obdobju 1. 2. 2021 – 31. 1. 2025 sodeluje v mednarodnem H2020 konzorcijskem projektu **Athena**: Izvajanje načrtov za enakost spolov za sprostitev raziskovalnega potenciala v raziskovalnih organizacijah in organizacijah za financiranje raziskav v Evropi
- leta 2021 ustanovljen **Odbor za izvedbo načrta za enakost spolov** (INES), ki ima pomembno vlogo pri ozaveščanju, spremljanju in uvajanju institucionalnih sprememb na IJS za uresničevanje enakosti spolov
- IJS ima **osebo za področje enakosti spolov, etičnih vidikov in reševanja sporov** v Direktorjevi pisarni, Ga. Iva Perhavec. Gre za prvo tovrstno zaposlitev osebe, ki na institucionalni ravni pokriva vse vidike uresničevanja enakosti spolov ter enakih možnosti žensk in moških ter aktivno sodeluje pri načrtovanju, razvoju in izvedbi ukrepov predvidenih po Načrtu za enakost spolov.
- **Načrt za enakost spolov** je bil sprejet oktobra 2022
- Cikel **internih izobraževanj** na področju enakosti spolov jeseni 2022
- Sedaj poteka izbor in določitev **ključnih indikatorjev** na področju enakosti spolov ter zbiranje kadrovskih podatkov, ločenih po spolu, priprava smernic za **spolno občutljivo rabo jezika** in komunikacijske aktivnosti

Hvala za pozornost

Napoved: Delavnica na temo AI&bias v Ljubljani, 16.5.

alenka.gucek@ijs.si

