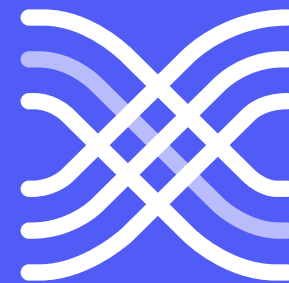


# Hands on session: case studies on detecting bias



# AI4Gov

Trusted AI for Transparent Public Governance  
fostering Democratic Values

**Alenka Guček**

alenka.gucek@ijs.si

Department for artificial intelligence

Jožef Stefan Institute

Ljubljana, Slovenia



# Group work

Discuss how bias can influence the algorithms being built for one of the 3 hypothetical scenarios.

Identify:

- at what stages the bias can occur,
- which types of biases can slip in the system and
- how the bias affects the users.

Timeline:

20 min discussion in a group

20 min report from groups and panel discussion



# Setting

Participants are grouped in groups of 3-5 people.  
Participants on Zoom are grouped in breakout rooms.

Make sure the groups are heterogenous, including someone who is a data scientist, someone from the policy making side, others.

Mentors that will help are assigned to the groups (formally or someone is informally going around and checking how is it going for them)



# Scenarios

Discuss how bias can influence the algorithms being built for one of the 3 hypothetical scenarios.

Identify at what stages the bias can occur, which types of biases can slip in the system and how the bias affects the users.

SCENARIO 1: social services automation of child care subsidised.

Aim: social service workers are overwhelmed with the amount of paperwork they have. The idea is to build an automated system, that will ingest the data users put in the system through a custom made user interface and build an algorithm that will decide eligibility and the amount of child care subsidised based on the historical data.

SCENARIO 2: recommender system on municipality website.

Aim: Municipality of a smaller city is trying to reach out to their inhabitants. They have updated their website (desktop and mobile friendly), and want to develop a recommender system that will offer news and information based on the history of browsing. Their idea is to classify people into seniors, adults, adults with young kids, and teenagers. They have hired a recent CS graduate to do the job.

SCENARIO 3: computer vision recognition on a metro.

Aim: lots of thefts are happening on the metro stations. City wants to prevent them by placing cameras on the metro stations and by utilising a private company-developed computer vision algorithm that can recognize the faces and match them with the official records. What could go wrong?