

Equality, artificial intelligence, algorithms and automated decisions:

how to detect and tackle discrimination
in the context of artificial intelligence?

Miroslav Chlipala



This training session is funded under the 'Rights, Equality and Citizenship Programme 2014-2020' of the European Commission.

Frequent use of the term

<>

different interpretation of the meaning

There is no official definition

The tradition of sci-fi

General artificial intelligence

vs.

Narrow artificial intelligence

General artificial intelligence

represents a machine that can handle any intellectual task

Narrow artificial intelligence

denotes a system that performs a single action

General artificial intelligence is the domain of science fiction

All current artificial intelligence methods include
into the realm of narrow artificial intelligence!

Key concepts

Autonomy

Ability to perform tasks in a complex environment without
constant user guidance

Adaptability

Ability to improve performance by learning from experience

The huge potential of artificial intelligence

Health
Transport
Energy
Agriculture
Tourism
Cybersecurity

Trusted artificial intelligence

Trust as a prerequisite
Ethical technology

European concept
trusted artificial intelligence

Regulation on artificial intelligence

Concept of risk

Unacceptable risk

High risk

Limited risk

Minimum risk

Unacceptable risk

Forbidden AI systems

Threats to the safety, life and rights of people

High risk

Systems with anti-discriminatory impact as well

e.g. education (test scoring)
staff management (sorting CVs)

High risk

Strict obligations:

- Risk assessment and mitigation
- High quality data sets entering the AI system
- Activity logging (log management)
- Detailed documentation
- Clear and adequate information for the user
- Adequate human supervision
- High level of reliability, safety and accuracy

High risk

- Remote biometric identification
- Exceptions (with the permission of a court, other independent authority)
 - Search for the missing child
 - Preventing a specific and imminent terrorist threat
 - Perpetrator of a serious crime

Limited risk

(e.g. chatbots)

User Knowledge
for his decision whether to use such a system

Minimum risk

Video games with AI
Spam filters

It poses only a minimal risk to rights and security
(Regulation does not apply to these systems)

Commission proposal

National supervisory authorities
European Council on Artificial Intelligence

Autonomous vehicles

(delivery robots, drones or autonomous vessels)

Recommended content

While the headline of the printed version of the "newspaper" is the same for all readers, the homepage of the online version is different for each reader

The algorithms that determine the content displayed are based on artificial intelligence

Image and video processing

(facial recognition - photo sorting, automatic tagging in social media)

Generating or editing digital content - photo filters

Computer-generated characters in films
(imitation of human gestures)

Thank you for your attention!

Your questions...?