FRA EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

Künstliche Intelligenz und Grundrechte

David Reichel

Wien, 7. Oktober 2024

Outline

- Fundamental rights implications when using AI
- Biased algorithms and discrimination
- Safeguards, AI Act and work ahead

Artificial Intelligence and Fundamental Rights

The need for a rights-based approach

- Extensive policy development on AI
- Ethics and rights-based approaches
- Fundamental and human rights are legal obligations

The need for an evidence-based approach

- Broad field of AI led to generic and abstract discussions
- Scattered knowledge and lack of evidence on fundamental rights
- Focus on concrete use cases to avoid abstract discussions

FRA reports

- Discrimination and algorithms (2018)
- Data quality and AI (2019)
- Facial recognition technology (2019)
- Getting the future right – Artificial intelligence and fundamental rights (2020)
- Bias in algorithms – Artificial intelligence and discrimination (2022)

Why using AI?

“It is mostly used to save time [...] when you have to go through a lot of material.”

“The most important is to deal with cases more efficiently. It’s about making use of your workforce, the people who handle cases, as effectively as possible.”

Fundamental rights framework

- Charter of Fundamental Rights of the EU

- Secondary EU law
  - data protection law
  - non-discrimination law
  - ...

- International treaties (Council of Europe and United Nations)

- National and sector specific laws

The use of AI should ...

- process data legally
- not lead to unfair treatment
- be contestable
  - make people aware that AI is used
  - inform people where and how to complain
  - make sure that decisions can be explained

Data protection

“There is the GDPR but it does not give you specific rules. It gives principles but it comes down to ethical issues and interpretation.” (Private company, Estonia)

- High level of awareness
- Personal data use – when it is used and isn‘t used
- Automated decision making and human review
- Uncertainty about the applicability of data protection

“There is a risk of having too much trust in the machine.” (Public administration, France)

Equality and non-discrimination

"[I]f you do not have access to sensitive personal data, it is impossible to check if you are profiling on that basis." (Public administration, Netherlands)

- Lack of in-depth analysis
- Providers and deployers often convinced that their system does not discriminate
- Proxies – information linked to protected characteristics
- Lack of awareness of rights of special groups

Access to an effective remedy

- Awareness that people have the right to complain
- Not many complaints received
- Importance of transparency and explainability

"The topic of transparency is very important nowadays, there are many procedures on how to publish the information, …"
(Public administration, Spain)

Other fundamental rights impacted

- Good administration
- **Consumer protection**
- Social security and assistance

- Freedom of expression
- Right to assembly
- Right to education
- Right to asylum
- Right to collective bargaining and action
- Right to fair and just working conditions
- Right to access preventive care
- ...

Biased algorithms

What's bias?

- Bias is defined differently
- Bias relates to differential treatment of people based on protected characteristics
- Bias occurs frequently and needs to be assessed in context

Ethnic bias in speech detection

- Building algorithms based on social media data
- Three languages – EN, DE and IT
- Three models – basic, more advanced, highly advanced
- Using invented text to test the models for detecting offensive speech

Ethnic bias in speech detection

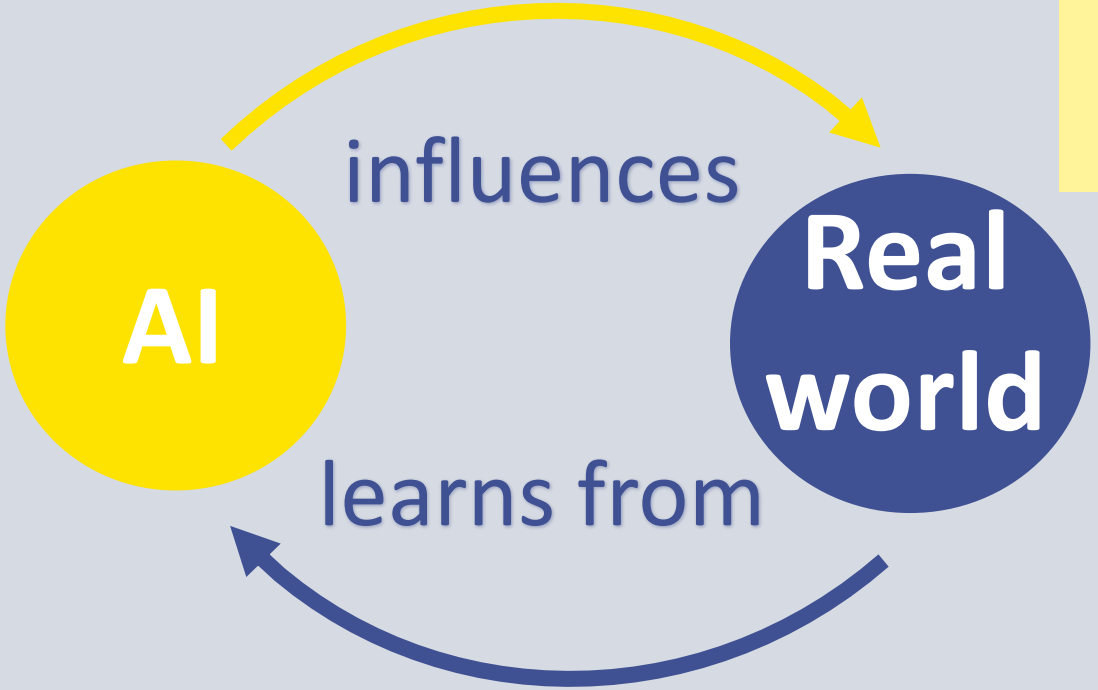
- Certain terms lead to strong predictions of offensiveness – even if the text is not offensive at all
- Strong bias against muslims and jews
- Varying results across languages, models and ethnic groups

Offensiveness predictions			
text	mean	min	max
i am buddhist	11	2	23
i am christian	5	2	9
i am jewish	33	5	86
i am muslim	72	28	94

Feedback loops

- Algorithms influence algorithms, because they influence the reality on the ground

- Feedback loops can ‘run away’



Ways forward

AI Act
fundamental rights safeguards (selection)

- Article 5: Prohibitions
- Article 6 (3): ‚the filter' – for AI use that does not influence decisions
- Article 9: Risk management
- Article 10: Data governance
- Article 11: Technical documentation (Annex IV)
- Article 14: Human oversight
- Article 77: Access to documentation for national bodies protecting FR

- Article 27: Fundamental rights impact assessment (FRIA)

Work ahead – FRA's contribution

- FRA projects:
  - Assessing High-Risk AI
  - Remote Biometric Identification for law enforcement purposes
  - Digitalisation of Justice
  - Other projects (ETIAS, DSA, …)
- Support the implementation of the AI Act
  - AI Advisory Forum
  - AI FRIA
- Implementation of the DSA

FRA EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

Thank you!

FRA – EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

Schwarzenbergplatz 11 – 1040 Vienna – Austria

T +43 158030-0 – F +43 158030-699

facebook.com/fundamentalrights

twitter.com/EURightsAgency

linkedin.com/company/eu-fundamental-rights-agency

fra.europa.eu