Text, logo

Description automatically generatedLogo

Description automatically generated with medium confidence

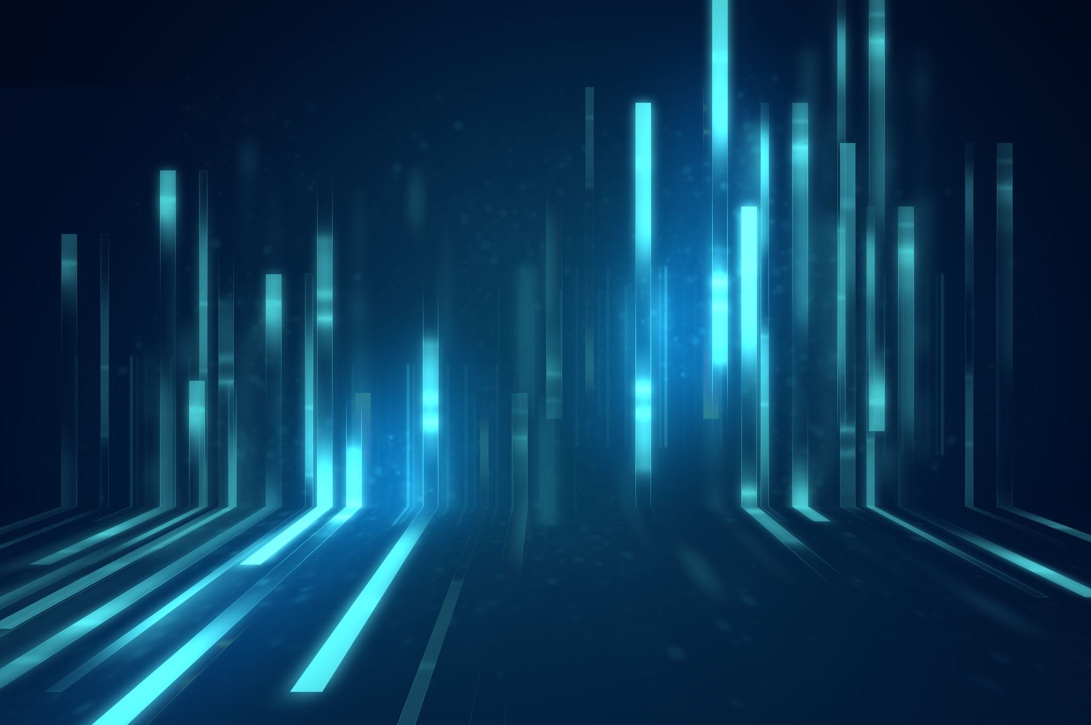
­­­

Fairness

Intermediate

Module 3

Legal Interpretations of Fairness and Transparency



# Introduction

This module will interface and expand on the UK's GDPR and the EU's AI Act. It will introduce a focus on soft law and practical use cases, including the obligations set out in the EU's AI proposed Act, soft law guidance from the authorities on fairness and transparency, UK standards for algorithmic transparency and the EC Guidelines on Trustworthy AI.

­­­

# Pre-reading

* Elon Musk: regulate AI to combat 'existential threat' before it's too late, The Guardian: <https://www.theguardian.com/technology/2017/jul/17/elon-musk-regulation-ai-combat-existential-threat-tesla-spacex-ceo>
* Cambridge Analytica scandal highlights need for AI regulation, The Guardian, April 2018: <https://www.theguardian.com/technology/2018/apr/16/cambridge-analytica-scandal-highlights-need-for-ai-regulation>
* Towards Intelligent Regulation of Artificial Intelligence, Cambridge University Press:<https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/towards-intelligent-regulation-of-artificial-intelligence/AF1AD1940B70DB88D2B24202EE933F1B>

# Reading

# China and Europe are leading the push to regulate AI – one of them could set the global playbook, CNBC, May 2022: <https://www.cnbc.com/2022/05/26/china-and-europe-are-leading-the-push-to-regulate-ai.html>

# UK National AI Strategy – AI Action Plan, Gov.Uk, Office for AI, July, 2022: <https://www.gov.uk/government/publications/national-ai-strategy-ai-action-plan/national-ai-strategy-ai-action-plan>

# A European approach to artificial intelligence, European Commission: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>

* First regulatory sandbox on Artificial Intelligence presented, European Commission, June 2022: <https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented>
* How to Train an AI with GDPR Limitations: <https://intellias.com/how-to-train-an-ai-with-gdpr-limitations/>



­­­

# Tasks

**Task 1 – Real world example**

In 2013, Dutch tax authorities launched a self-learning algorithm to create risk profiles in an effort to spot childcare benefits fraud. Between 2013 and 2019, thousands of persons were wrongly accused of committed fraud and were forced to pay large amounts of money. In many cases, this sum amounted to tens of thousands of euros, driving families into severe financial hardship. As this case shows, mistakes in the implementation of machine learning, can lead to unfair outcomes. The data was sometimes wrong and there was a late involvement of the privacy officer and the system showed unfair bias towards certain groups of the population, for instance those with double nationality and immigrants.

In April 2022, the Dutch Data Protection Authority fined the Ministry of Finance with EUR 3.7M for violations of the GDPR.

QUESTION: what measures could have been implemented to avoid the harm caused on the individuals? To be discussed in groups.

**Task 2 – Open Questions**

What measures should be implemented to ensure fairness and transparency in the following AI systems according to what was learnt during the lesson? Open discussion

- An airline uses a chatbot operated with AI as part of their customer support online

- An airport uses an AI system to check, amongst others, the identity of travellers. As a result, travellers will be allowed to cross the border or refused entry.

- A company uses a software operated with AI to facilitate the recruitment process, the system automatically checks and reads the CVs received and decides whether to invite applicants for an interview or not.

- The police are seeking to implement AI for real-time remote identification under certain circumstances.

**Self-assessment pass/fail questions**

1. When was the Proposal for an AI Regulation in the EU published?
2. In 2020
3. In 2021
4. In 2022
5. When will the AI proposed Regulation come into force?
6. In 2023
7. In 2024
8. It is not yet known
9. Will the AI proposed Regulation apply to non-EU providers of AI?
10. In some instances
11. Always
12. Never
13. What does the risk-based approach of the AI proposed Regulation entail?
14. Different obligations depending on the risk level
15. Different obligations depending on the sector
16. Different obligations depending on the type of AI
17. How is the real-time biometric identification classified?
18. As an unacceptable risk
19. High-risk
20. Limited risk
21. How will AI systems intended to be used as safety components in road traffic management will be classified?
22. As an unacceptable risk
23. High-risk
24. Limited risk
25. Which of the following is a transparency obligation for high-risk AI systems?
26. Interpretation of the system’s outputs
27. The inclusion of instructions for use
28. Both (a) and (b) are correct
29. Which of the following are transparency obligations for minimal risk systems?
30. Human oversight measures
31. Provision of information about the provider
32. Information about the use of deep-fakes
33. Which of the following are measures to ensure fairness?
34. Data governance
35. Record keeping
36. Both (a) and (b) are correct
37. If the AI system process personal data, when should privacy teams be engaged and consulted?
38. At an the early stage in the design process
39. When the AI system is being trained
40. When the AI system is being validated and tested
41. How many principles are enshrined within the ‘Ethics Guidelines for Trustworthy AI’?
42. 5
43. 7
44. 6
45. Who is subject to the UK Algorithmic Transparency Standard?
46. UK public sector entities
47. UK private sector entities
48. Both (a) and (b) are correct
49. A US company is operating an AI system in Spain which process personal data of its French customers, will the GDPR be applicable?
50. Yes
51. No
52. It depends
53. Which of the following are transparency obligations under the GDPR?
54. The provision of information about international data transfers
55. The existed of automated-decision making and meaningful information about the logic involved.
56. Both (a) and (b) are correct

**Answers**

# Answers to tasks

**Task 1**

- Human intervention or human in the loop measures

- Involving the privacy officer in the early stages of the project

- Maintaining the data accurate: accuracy of the data and the datasets

- Implementation of the system monitoring measures

- Checks against unfair bias

Final remark: In the aftermath of the Dutch Scandal, the European Committee of the Regions (https://cor.europa.eu/EN/our-work/Pages/OpinionTimeline.aspx?opId=CDR-2682-2021) proposed to introduce the necessity for human intervention in high-risk AI systems intended to be used by public authorities to evaluate the eligibility of natural persons for public benefits and to evaluate creditworthiness. The inclusion of this requirement is still under discussion.