AI and Policymaking Workshop

9th April 2024

Co-funded by
the European Union

www.aiskills.eu

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

ARISA – Artificial Intelligence Skills Alliance

- ARISA delivers a strategic approach to sectoral cooperation on the development of **Artificial Intelligence (AI) skills** in Europe. The project provides AI knowledge and skills helping people understand and use AI technology in business and policy contexts, taking into account privacy, bias, and trust.

- ARISA is a four-year transnational project funded under the Erasmus+ programme. The project has 18 partners and 28 associated partners.

2022/2023

Needs Analysis and a European Strategy for AI skills development

2023/2024

AI skills curricula & learning programmes, certification methods & framework

2024/2025+

Learning programmes & courses piloting and further uptake of the ARISA results

ARISA fast-tracks the upskilling and reskilling of employees, job seekers, business leaders, and policymakers into AI-related professions to open Europe to new business opportunities.

Today’s speakers

Julien Chasserieau

Associate Director for AI and Data Policy, DIGITALEUROPE

Julien analyses AI and data economy policies at EU level and leads advocacy campaigns for the tech industry. He has worked on AI policy since 2019.

He works with a wide range of industry stakeholders as well as institutional and academia partners, including the 110 companies and 40 national trade associations members of DIGITALEUROPE.

Joris Krijger

AI & Ethics Officer, de Volksbank

Joris is an Ethics & AI Officer at de Volksbank while also holding a PhD at the Erasmus University in Rotterdam, focusing on AI ethics.

He presently works on bridging the gap between ethical principles and AI by studying the operationalisation of ethics in data science contexts. He is part of the editorial board of Springer Nature’s AI and Ethics Journal. For his contributions to the field of responsible AI he received a Graduate School Award for PhD Excellence.

Cornelia Kutterer

Managing Director, Considerati & Senior Research Fellow at the Multidisciplinary Institute in Artificial Intelligence (Chair of AI Regulation), University of Grenoble-Alpes

Cornelia leads Considerati in Belgium, provides legal and AI governance services, and researches AI regulation as a Senior Research Fellow at the University of Grenoble. She is a tech law expert, IAPP board member, and part of the OECD AI network. She advises Safer.AI and was a lead at Microsoft Europe in tech policy. Her background includes roles in European consumer law and early experience in legal institutions. She holds law degrees from Hamburg and Strathclyde Universities.

Agenda

13:30 – 13:35	**Welcome & Introduction** *Jose Martinez-Usero, Projects Director & Niels Selling, Senior Research & Innovation Manager, DIGITALEUROPE*
13:35 – 14:35	**European AI policies and their impact on AI Use and Innovation** *Julien Chasserieau, Associate Director for AI and Data Policy, DIGITALEUROPE*
14:35 – 14:45	Coffee break
14:45 – 15:45	**AI Governance: Ensuring Responsible Deployment** *Joris Krijger, AI & Ethics Officer, de Volksbank*
15:45 – 16:00	Coffee break
16:00 – 17:00	**Charting the Future – Exploring compliance readiness, safety innovation and open legal questions** *Cornelia Kutterer, Managing Director, Considerati*

European AI policies: impact on AI use and innovation

Brussels, 9 April / Julien Chasserieau, DIGITALEUROPE

Co-funded by the European Union

www.aiskills.eu

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

AI initiatives at a glance

European initiatives

AI Act

Product Liability Directive / AI Liability Directive

New Legislative Framework
Product safety laws

EU **AI innovation strategy**

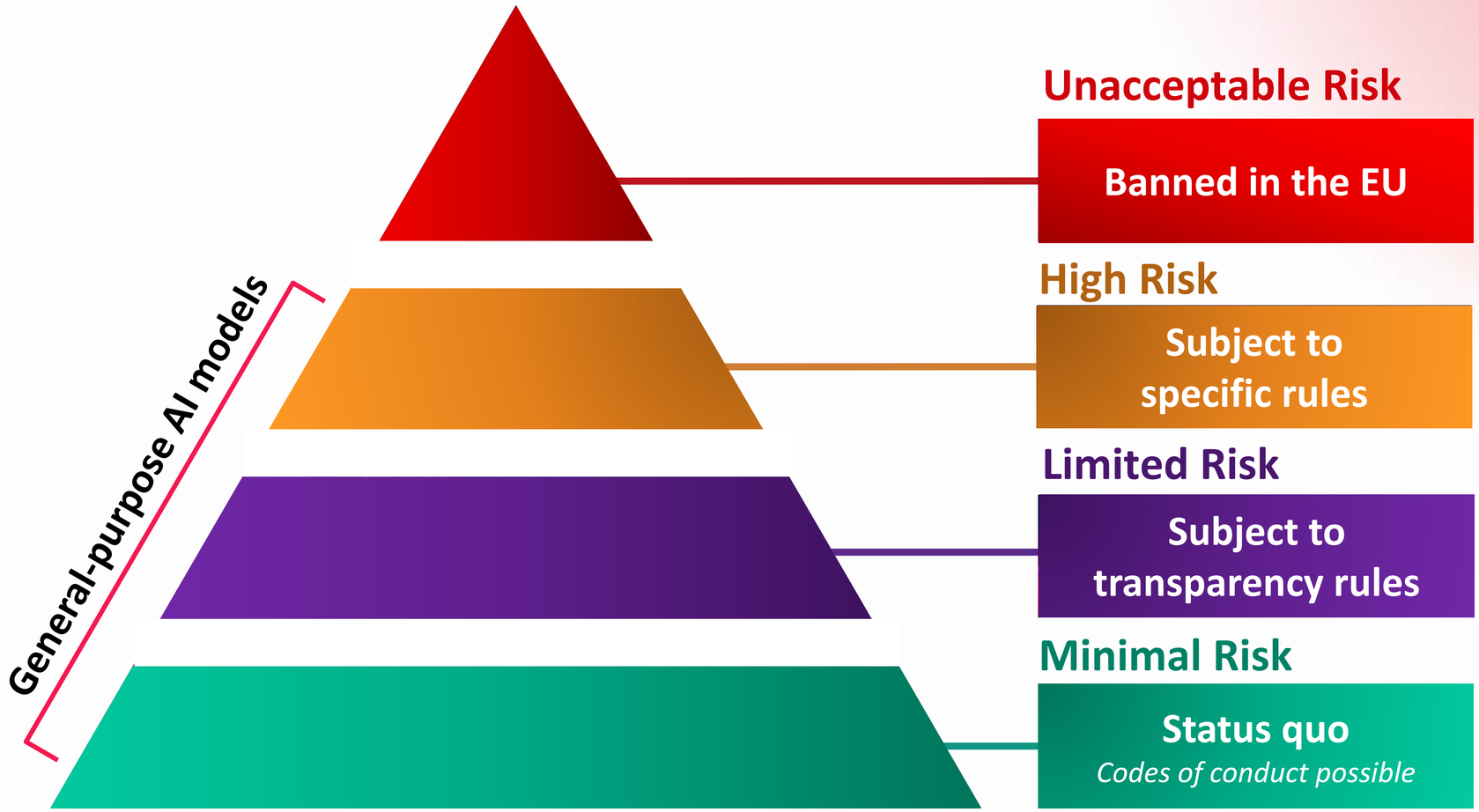
Global initiatives

Council of Europe Framework on AI, Human Rights, Democracy and the Rule of Law

G7 Hiroshima process
(for GPAI models)

UN AI Advisory Body Report

The AI Act risk-based approach



- Social scoring
- Mass surveillance
- *etc.*

- Recruitment
- Access to credit, insurance
- *etc.*

- Generative AI
- Emotion recognition
- *etc.*

arisa

AI definition

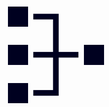
What makes a software an *AI system*?

A **machine-based** system

Designed to operate with **varying levels of autonomy,**

That **may exhibit adaptiveness after deployment**

And that, for **explicit or implicit objectives,**



Infers, from input, how to generate outputs – predictions, content, recommendations or decisions

That can **influence physical or virtual environments**.

Key exemptions from scope

National security and defence

AI used for scientific research

Free and open-source software

&

Member States can keep and create laws more favourable to workers.

Transparency obligations still apply.

Unless placed or put into service as forbidden or high-risk AI system.

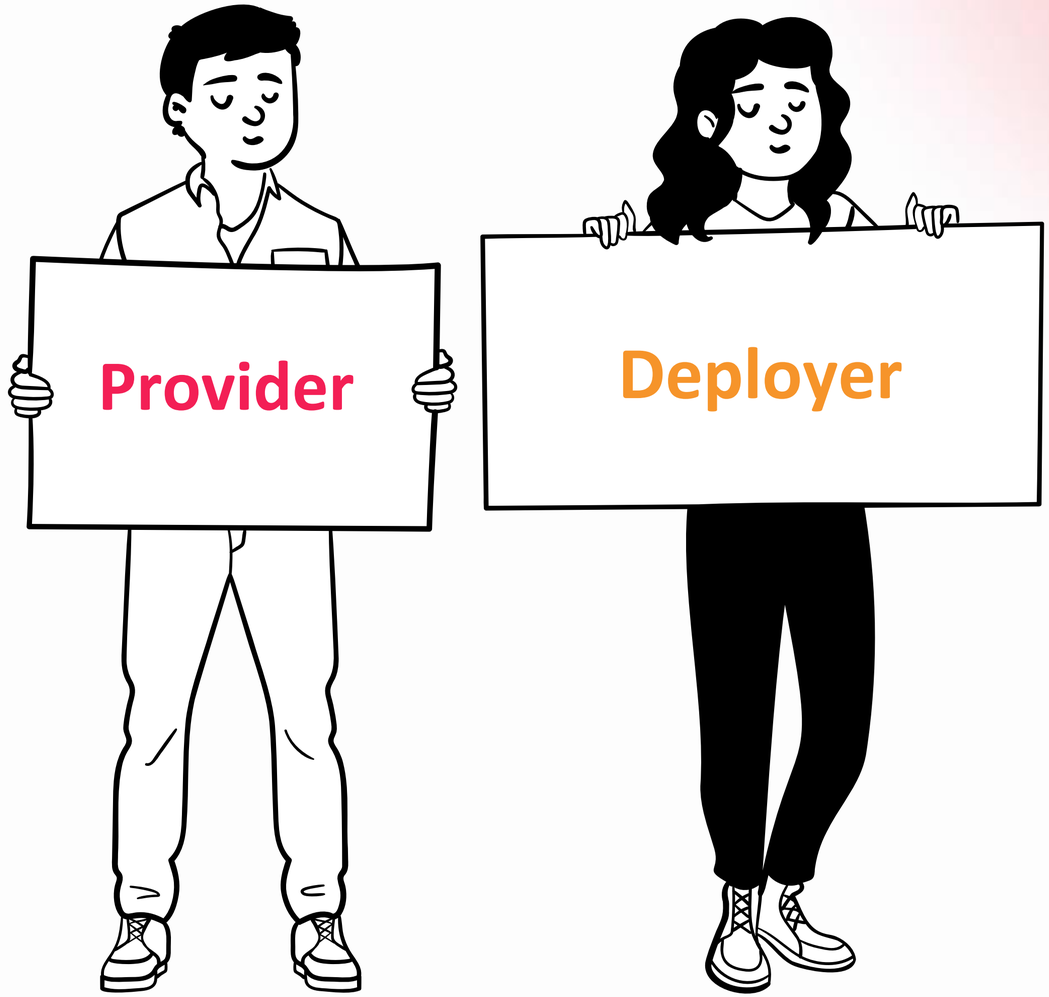
Providers vs deployers

A **person, company or organisation…**

That **develops** an AI system or GPAI model

Or has them developed and **places them on the market**

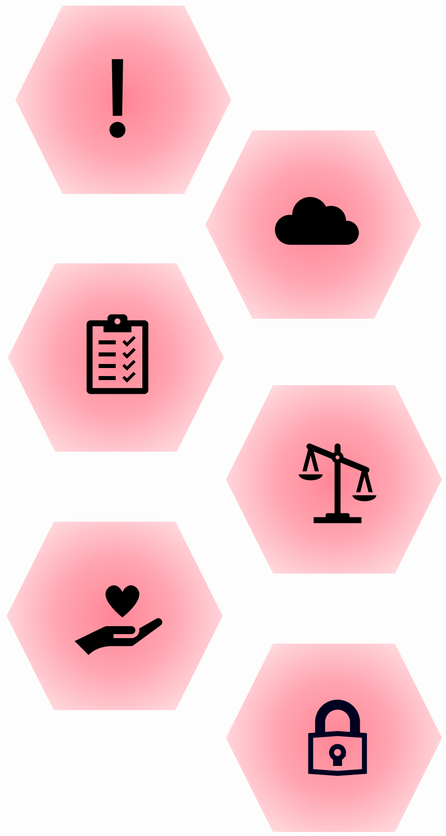
Or **puts the AI system into service** under its own name or trademark.



A **person, company or organisation…**

That **uses an AI system under its authority**.

Requirements for high-risk AI



- Risk Management System
- Data Governance
- Technical Documentation
- Record Keeping
- Transparency
- Human Oversight
- Robustness, Accuracy & Cybersecurity

Depending on risk level, GPAI models have different additional requirements

Additional transparency rules for AI that directly interacts with people or artificially manipulate/generate content

Obligations for providers & deployers of high-risk AI systems

Providers

Deployers

- Conformity assessments
- Share necessary info and provide technical access to deployers
- Inform & cooperate with competent authorities
- Develop a quality management system
- Develop post-market monitoring systems
- Report serious incidents & malfunctioning
- Feed relevant input, monitor operations & keep logs

Obligations for deployers of high-risk AI systems who are public authorities

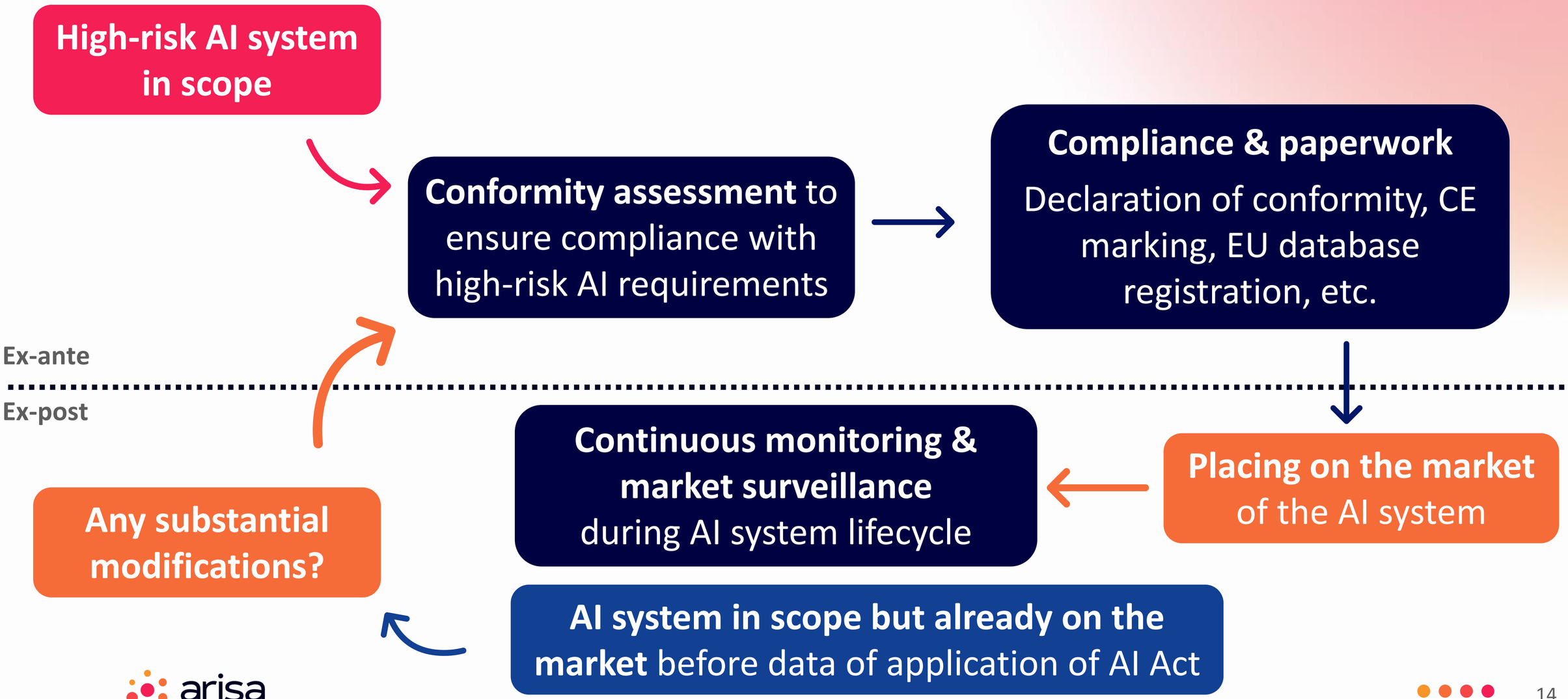
1. Public authorities, EU institutions, and those acting on their behalf, must **register themselves and the use of the AI system in an EU database**.

2. Before deployment, public authorities and entities providing public services (banks, insurance), must **conduct a Fundamental Rights Impact assessment, including a description of:**

Then **notify the market surveillance authorities of results via template**, provided by AI Office.

- **Processes and use of the AI system**
- **Time period of usage and frequency**
- **Affected natural persons**
- **Likely specific harms**
- **Human oversight measures**
- **Risk mitigation measures** (including internal governance)

AI Act compliance cycle



AI Act standardisation

Using **standards = presumption of conformity** with AI Act requirements. This means **easier compliance for organisations**.

European Commission sent standardisation request to CEN-CENELEC. Within CEN-CENELEC JTC 21, bottom-up process, led by experts within national standardisation bodies.

Using international standards but complementing with home-grown ones for value-based norms.

Deadline for deliverables: 30 April 2025 (likely delayed).

Measures to support innovation

AI regulatory sandboxes

- **To facilitate development and testing of AI** under regulatory oversight.
- + to **share best practices** and **make regulatory learning**.
- Operational **24 months after entry into force**.
- Provides **written proof of activities undertaken** and **exit report** which can be used to demonstrate compliance (≠ presumption of conformity).

Support to small providers & users

- **Priority** access to sandboxes.
- **Fee reduction** for 3rd-party conformity assessments.
- Potential **dedicated contact channels** to answer questions on implementation.

Sandboxes are mandatory but can be established jointly with other Member States.

Sandboxes in practice

Spain introduces first AI Regulatory Sandbox

- Cooperation between AI systems providers and deployers
- Test against requirements of the AI Act

Budget of €4.3 million over three years

- Funding from EU Recovery and Resilience Funds
- Linkin with Spanish National AI Strategy

Leading role in promoting innovation in AI

- Reporting of best practices, etc.
- Guidelines for European Commission
- AI advisory committee

Countries to watch

Denmark – Regulatory Sandbox
- Budget of €2.2 million (2024-2027)
- Judicial Counseling for AI Act & GDPR
- Deadline for application: 21 May 2024

Belgium – Sandbox Vlaanderen
- Not exclusive to AI – allows tests and experimentations

France – CNIL (data protection authority)
- Pilot AI sandbox: 8 projects to be tested
- Follows EdTech & digital health sandboxes

National implementation *for high-risk AI systems*

For GPAI models, enforcement is at EU level via AI Office

Member states must **designate as national competent authorities** *at least one*:

Notifying authority to select conformity assessment bodies to test compliance with the rules *before the AI is used*.

Market surveillance authority to test compliance *during the AI lifecycle*. If several authorities, one single point of contact must be chosen.

Need to be **independent** and have ***adequate*** **technical, financial & human resources**.

Other implementation activities

Build **regulatory sandboxes**

Coordinate with other **horizontal & sectoral authorities**

Monitor and report status of implementation to Commission

Appoint representatives to AI Board and Commission expert group

Getting involved

- **Support standardisation process** via national bodies

- **Provide input** on upcoming guidelines developed with AI Office
  - *Codes of practice*
  - *Codes of conduct*
  - *Implementing / Delegated Acts*

- **Provide education & training** – internally and externally

- **Promoting Research & Development** of trustworthy AI

- **Document and classify** internal use of AI

AI Act provisional timeline

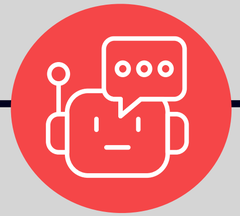
MID-2024

AI Act enters into force

EARLY 2025

Bans apply

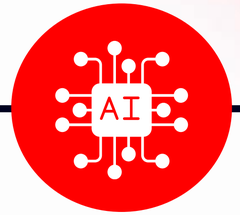
MID-2025



GPAI models rules apply

National authorities designated

MID-2026



High-risk rules apply for Annex III

Use categories:

Education, justice, biometrics, access to public services, etc.

MID-2027

High-risk rules apply for Annex I (NLF)

Product safety laws:

Toys, machinery, medical devices, etc.

arisa

Impact on the ecosystem

More trust in AI

Higher AI uptake
by citizens and consumers

AI Act can act as trust “label”
to stand out from competitors

Resource allocation

High investment needed
from public sector

Member States will need to invest
resources in skills, training staff and
capability building

Regulatory burden

High compliance costs
and paperwork

Compliance could cost €1.5 billion
to the EU economy in 2025*

*supporting study to the AI Act impact assessment (2021, Commission)

Compliance with the AI Act could cost around

€ 300 000

for an SME selling an AI system in the EU

Based on 2021 estimate from the Commission-supported study for the AI Act impact assessment, for a 50-employee SME, during the first year of compliance.

Liability framework for AI

Directives need be transposed by Member States into national law.

Product Liability Directive updated, to apply from end 2026.

- Directive sets EU-harmonised producer liability for damage caused by defective products.
- Scope includes digital manufacturing files and software, **including AI**.
- Compensation covers material and non-material losses.

Ongoing discussions on need for additional **AI Liability Directive**, specifically for AI systems, with link to AI Act.

- **2024 EU strategy to boost AI startups and innovation**

€3 billion public funding to harness potential of LLMs and genAI

Financial instruments for European AI startups

Support genAI talent pool with financial tools

Launch GenAI4EU strategy to stimulate genAI uptake

Support to European LLMs with high-quality data in all EU languages

“AI factories”: computing for AI model development

AI & robotics strategy expected in 2025

Q&A session

Feel free to take the floor!

Please use the “raise hand” feature

Responsible AI in Practice

Joris Krijger
Ethics & AI Officer | De Volksbank
PhD Ethics & AI | Erasmus University Rotterdam

Co-funded by
the European Union

www.aiskills.eu

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Agenda

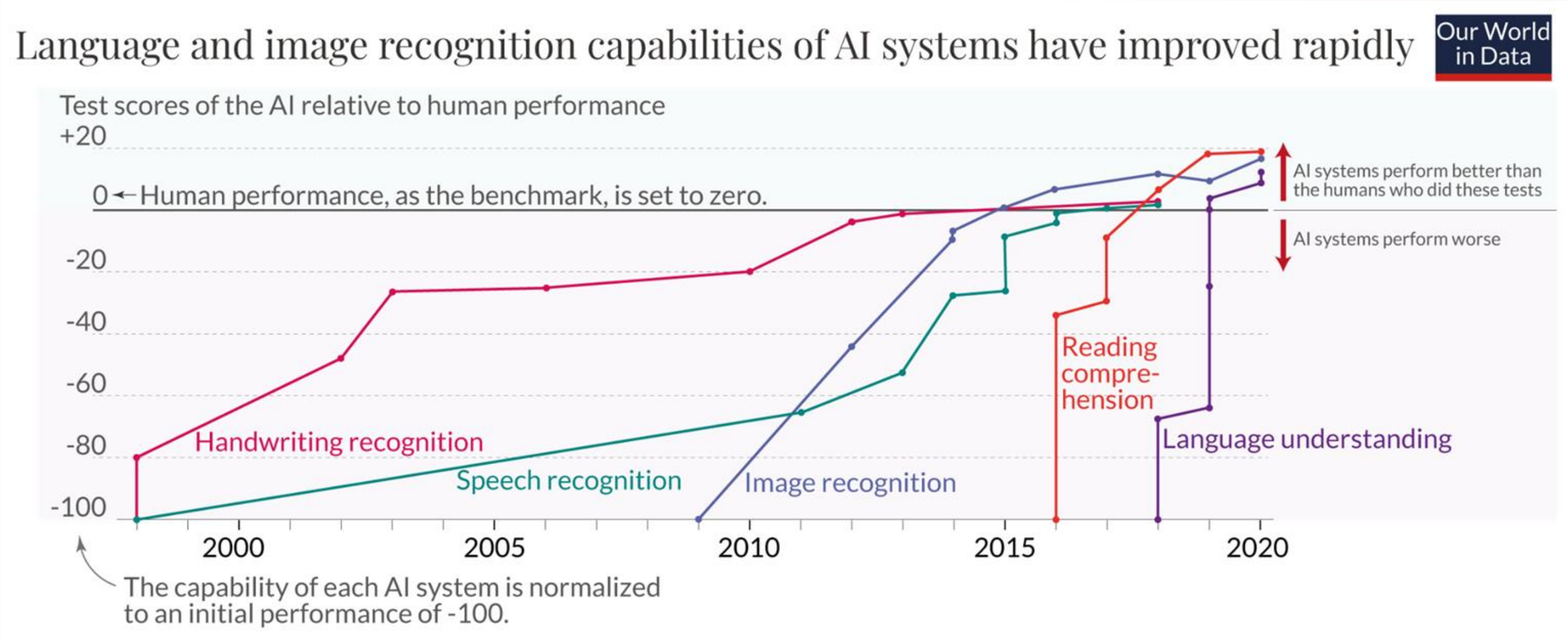
1. Understanding Responsible AI

2. Challenges for Current Frameworks

3. AI Ethics as Organizational Challenge

Understanding Responsible AI

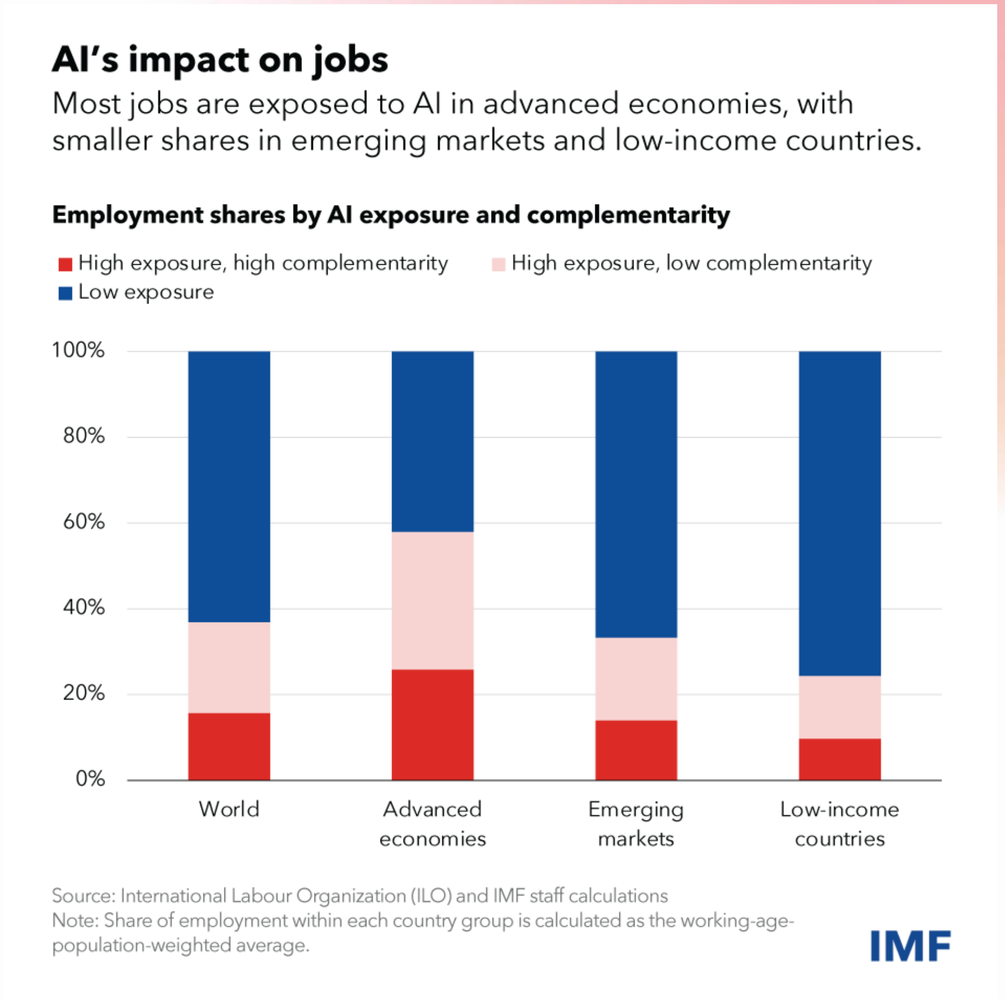
- **AI Impact**


Language and image recognition capabilities of AI systems have improved rapidly
Our World in Data
Test scores of the AI relative to human performance
+20
0 ← Human performance, as the benchmark, is set to zero.
-20
-40
-60
-80
-100
AI systems perform better than the humans who did these tests
AI systems perform worse
Handwriting recognition
Speech recognition
Image recognition
Reading compre-hension
Language understanding
2000
2005
2010
2015
2020
The capability of each AI system is normalized to an initial performance of -100.

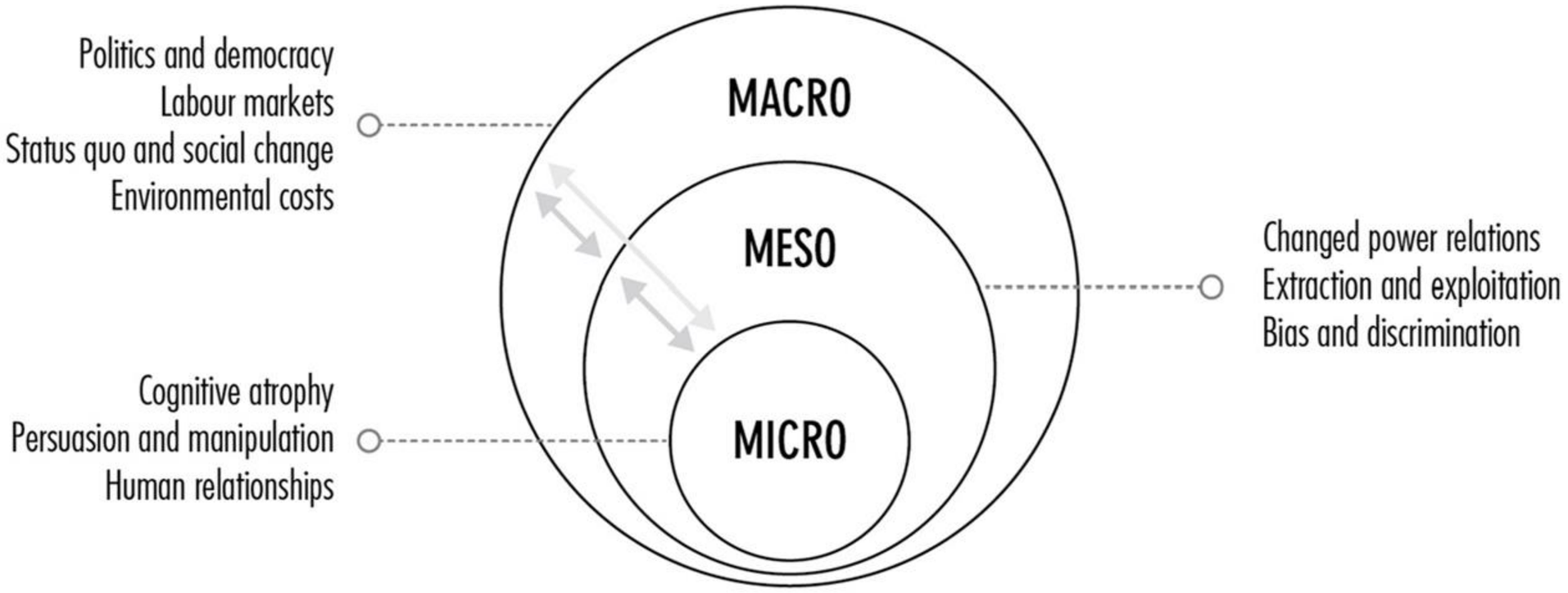
AI Impact

In advanced economies 60% of jobs will be affected by AI.

Roughly 30% will be complimentary, enhancing productivity. The other 30% will be automation, ultimately replacing jobs humans are currently doing.

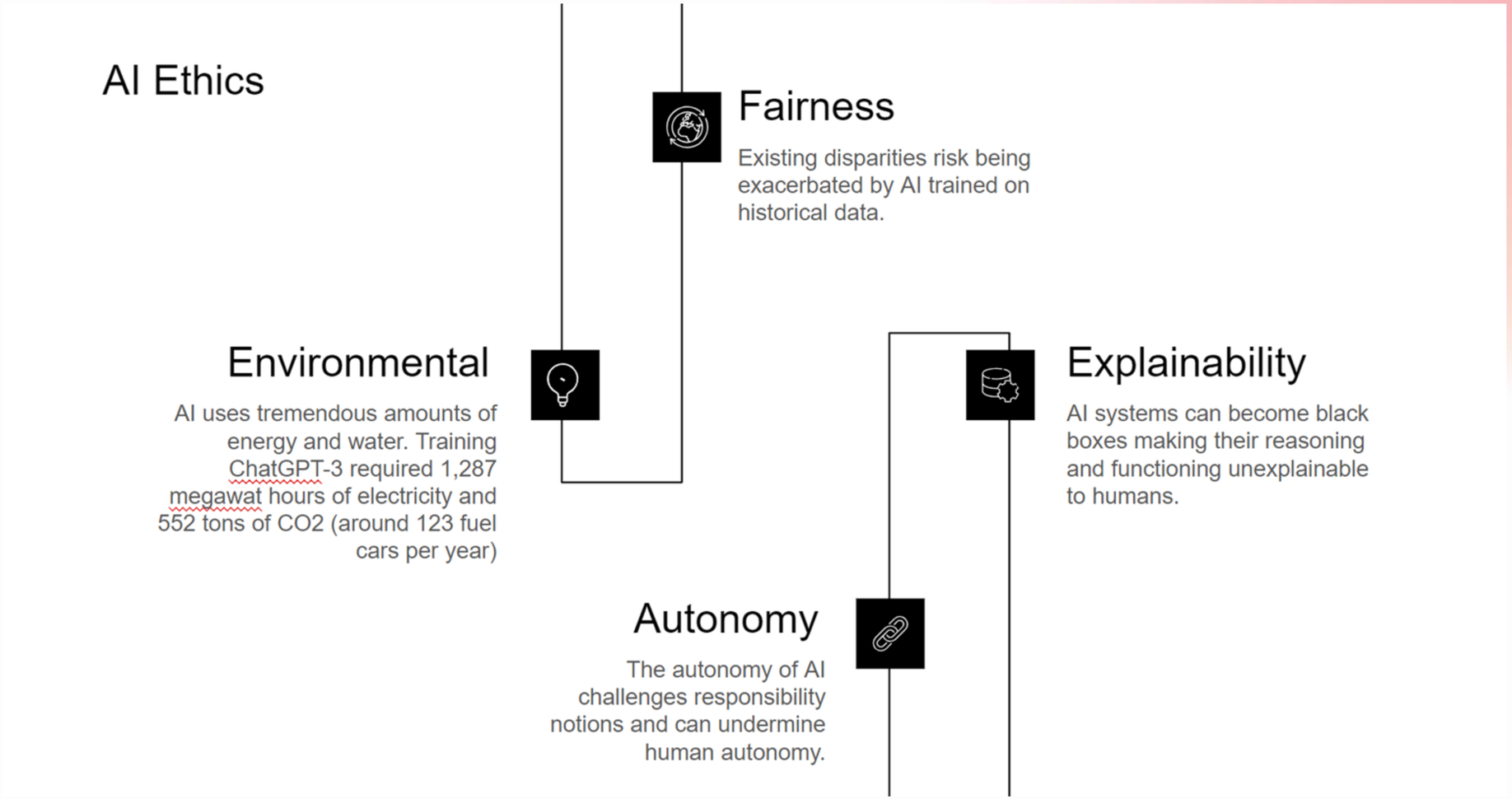


Ethics



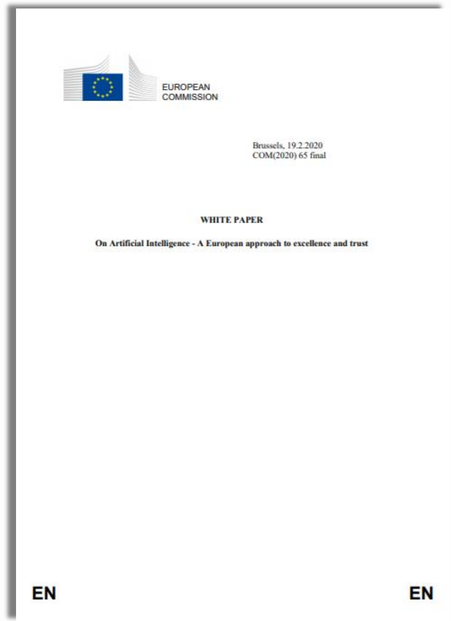
Henrik Skaug Sætr, 2023

Ethics

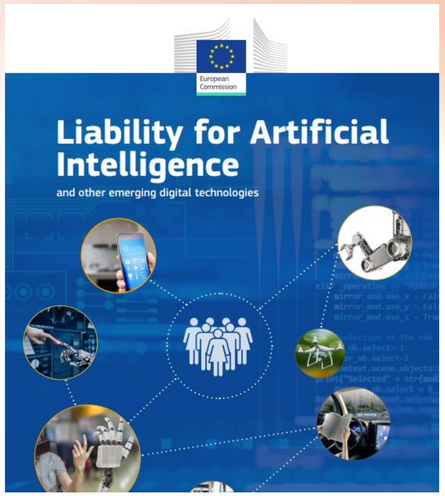

AI Ethics
Fairness
Existing disparities risk being exacerbated by AI trained on historical data.
Environmental
AI uses tremendous amounts of energy and water. Training ChatGPT-3 required 1,287 megawat hours of electricity and 552 tons of CO2 (around 123 fuel cars per year)
Explainability
AI systems can become black boxes making their reasoning and functioning unexplainable to humans.
Autonomy
The autonomy of AI challenges responsibility notions and can undermine human autonomy.

- **Frameworks**

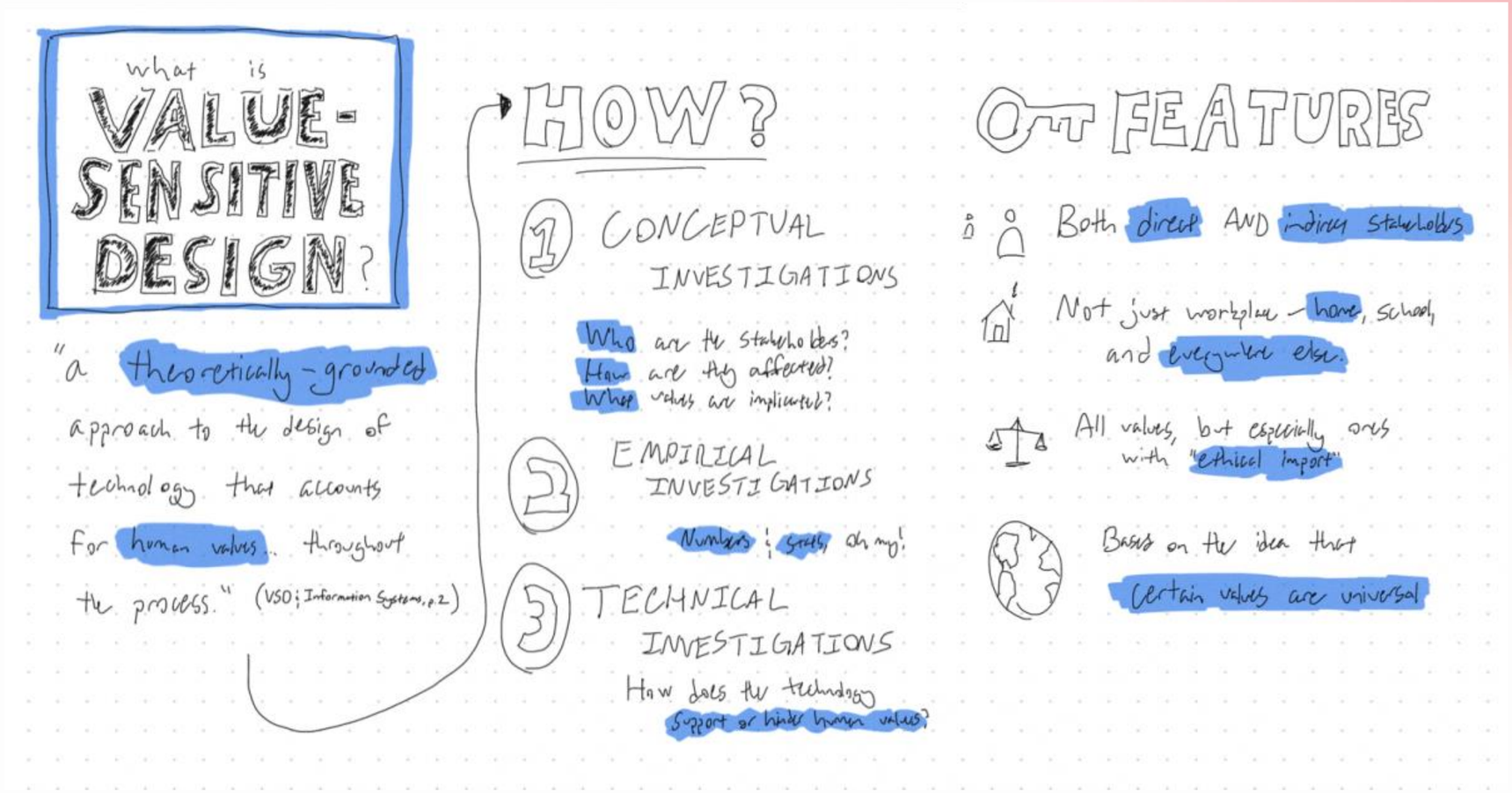

European Commission
INDEPENDENT
HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE
SET UP BY THE EUROPEAN COMMISSION
AI
ETHICS GUIDELINES FOR TRUSTWORTHY AI


EUROPEAN COMMISSION
Brussels, 19.2.2020
COM(2020) 65 final
WHITE PAPER
On Artificial Intelligence - A European approach to excellence and trust
EN
EN

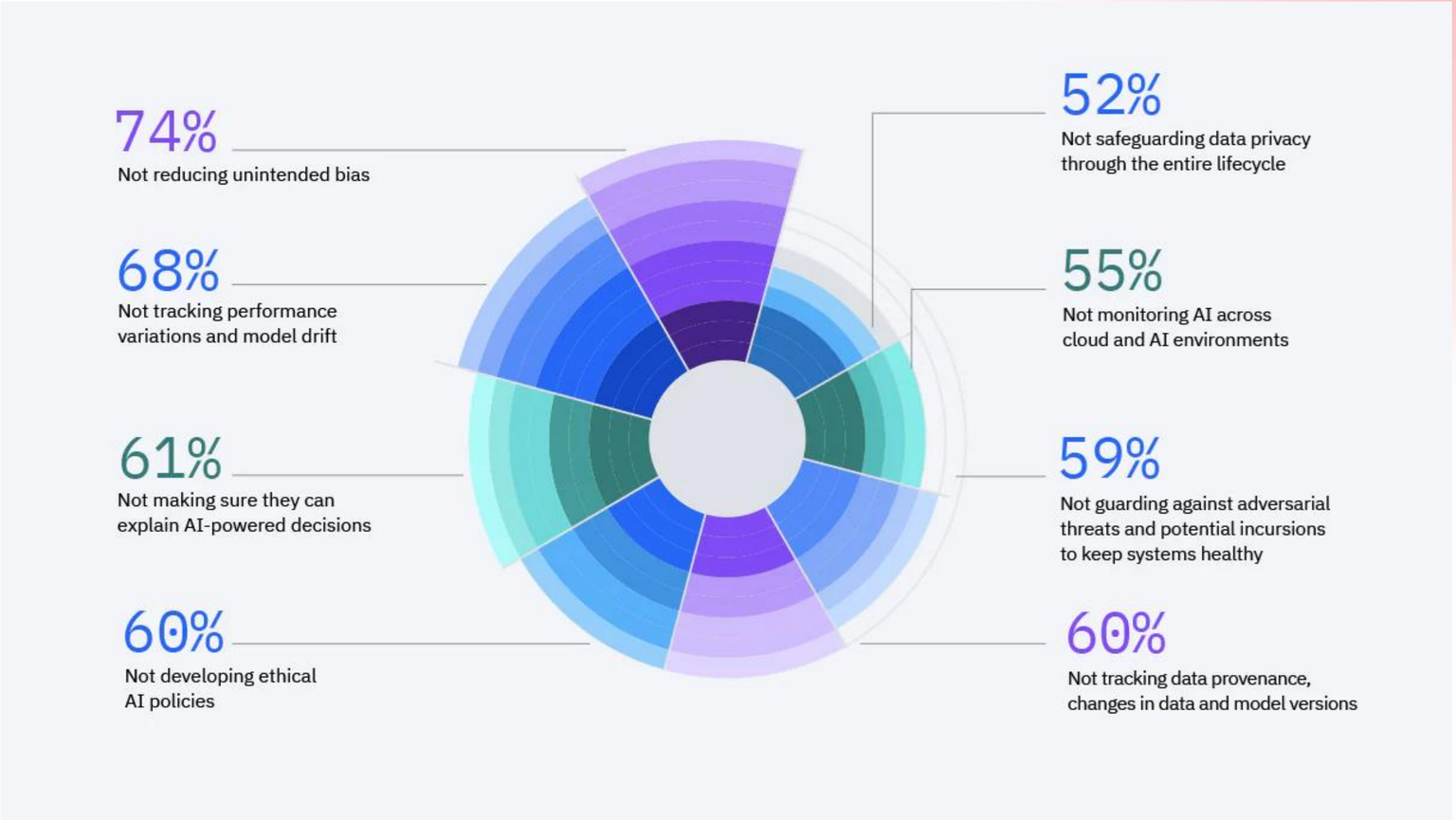

EU
Artificial
Intelligence Act


European Commission
Liability for Artificial Intelligence
and other emerging digital technologies

- **VSD**


what is
VALUE-SENSITIVE DESIGN?
"a theoretically-grounded approach to the design of technology that accounts for human values... throughout the process." (VSD; Information Systems, p.2)
HOW?
1 CONCEPTUAL INVESTIGATIONS
Who are the stakeholders?
How are they affected?
What values are implicated?
2 EMPIRICAL INVESTIGATIONS
Numbers; stats, oh my!
3 TECHNICAL INVESTIGATIONS
How does the technology support or hinder human values?
FEATURES
Both direct AND indirect stakeholders
Not just workplace - home, school, and everywhere else.
All values, but especially ones with "ethical import"
Based on the idea that certain values are universal

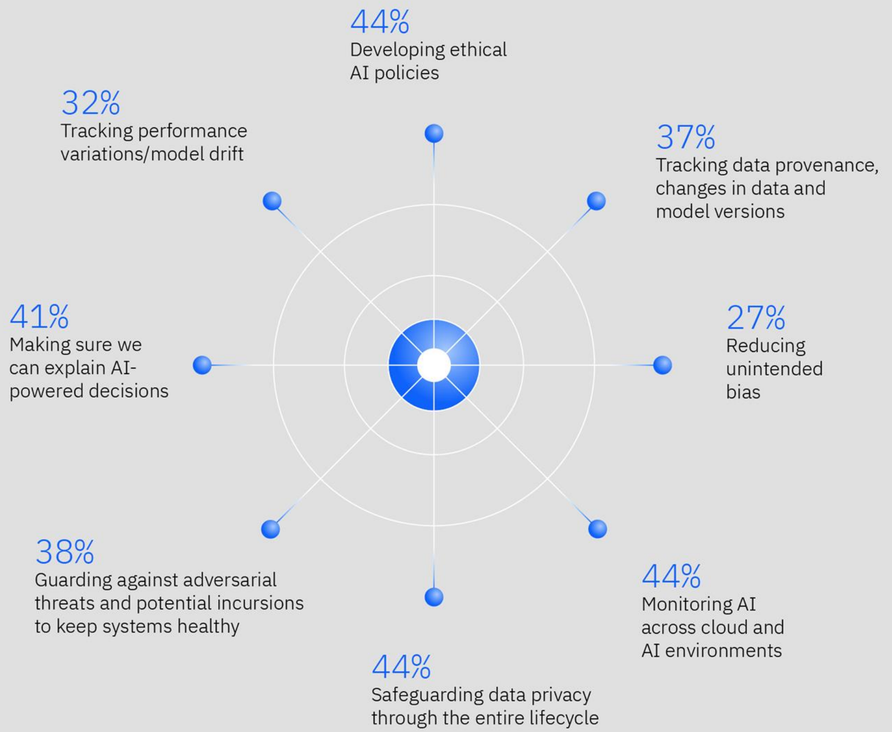
RAI



IBM AI Adoption Index 2022

RAI

Enterprises are taking steps to build trustworthy AI, but more progress is needed



IBM AI Adoption Index 20223

Challenges for Current Frameworks

AI Ethics Challenges

- Limited empirical success (McNamara, Smith and Murphy-Hill, 2018)
- Lack of accountability mechanisms, common practices and clear definitions (Mittelstadt, 2019)
- Malleable to actor agenda (Rudschies et al. 2021)
- Depoliticized and reductive by focussing on specific cases that are often underpinned by grander societal dilemmas
- Value conflicts

- Principles are unresponsive to context
- VSD is hard to consistently implement
- Delegation of responsibility to the individual level instead of questioning governance and power structures.

- Exaggerating the role and influence of ethics?

  "*[Guidelines] incorrectly portray the reasons why unethical technologies are found used in the wild today, making them appear like bone fide oversights that ethicists might be able to highlight rather than intrinsic parts of business models which disregard their effects on societies and environments.*" (Veale, 2020)

- **Value Conflict**

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

"We think ProPublica's report was based on **faulty statistics and data analysis**, and that the report failed to show that the COMPAS itself is racially biased, let alone that other risk instruments are biased. -- Flores et al. (2016)

- **Value Conflict**

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

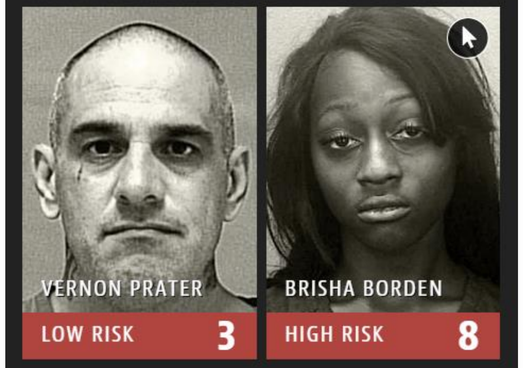
May 23, 2016

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

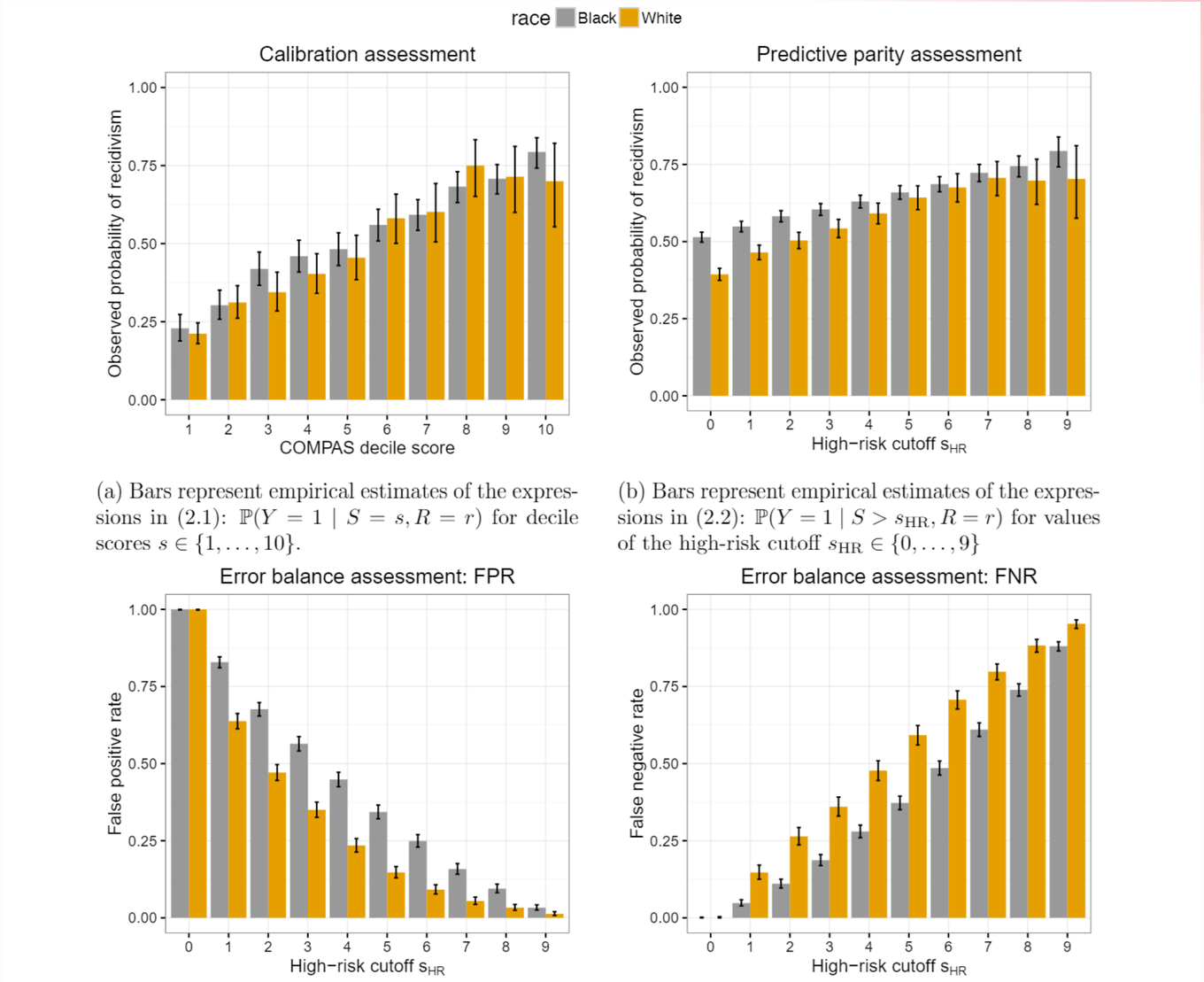
Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. *(Source: ProPublica analysis of data from Broward County, Fla.)*

Two Petty Theft Arrests



Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

Compas

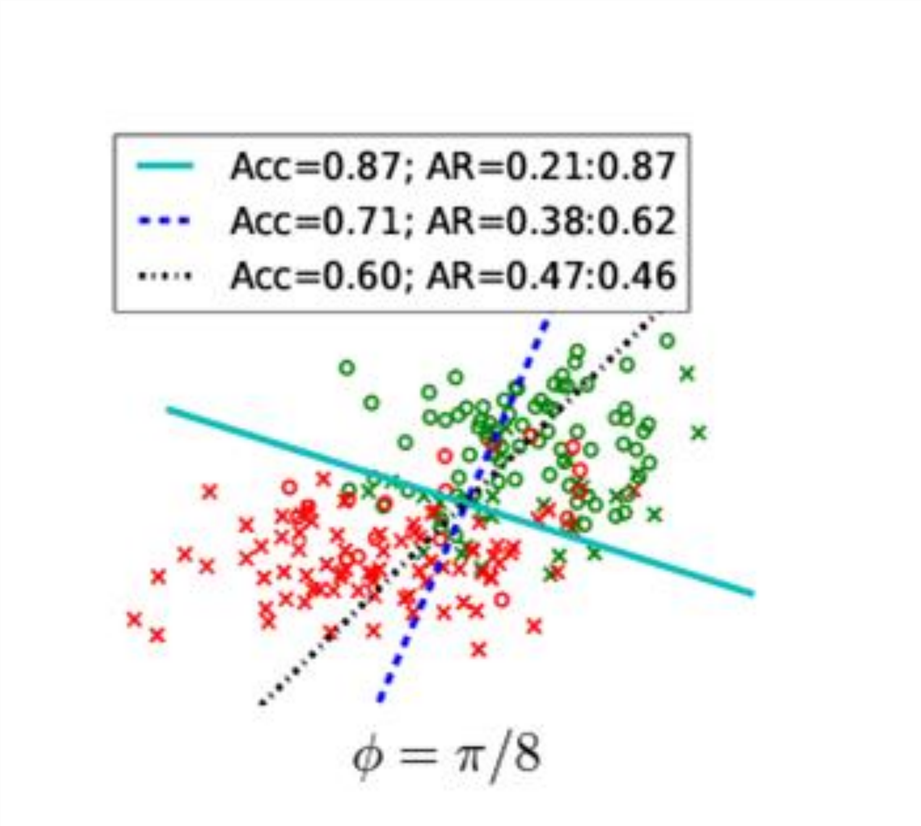


(a) Bars represent empirical estimates of the expressions in (2.1): $\mathbb{P}(Y = 1 \mid S = s, R = r)$ for decile scores $s \in \{1, \ldots, 10\}$.

(b) Bars represent empirical estimates of the expressions in (2.2): $\mathbb{P}(Y = 1 \mid S > s_{\mathrm{HR}}, R = r)$ for values of the high-risk cutoff $s_{\mathrm{HR}} \in \{0, \ldots, 9\}$

Value Conflicts I: Operationalization

Value Conflict II: Interactions



III. Interactions between Aspects

We describe and discuss the following interactions, based on support available in the literature:

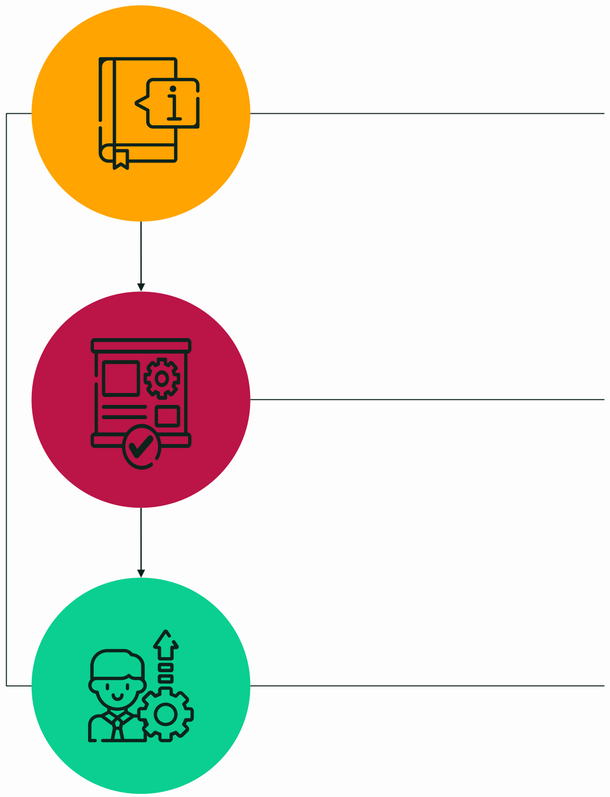
(a) Accuracy vs. Robustness
(b) Accuracy vs. Fairness
(c) Accuracy vs. Privacy
(d) Accuracy vs. Explainability
(e) Fairness vs. Robustness
(f) Fairness vs. Privacy
(g) Fairness vs. Transparency
(h) Privacy vs. Robustness
(i) Transparency vs. Explainability
(j) Transparency vs. Privacy and Robustness

The major challenges in responsible AI will not be technical but ethical in nature, requiring an ethical institutional infrastructure.

Joris Krijger

AI Ethics as Organizational Challenge

Principles to Practice?



Principles	High level principles
Organization	Organizational structures, culture and processes
Practice	Value sensitive design of specific AI applications

Three Pillars of Value Operationalization

Ethical Codes

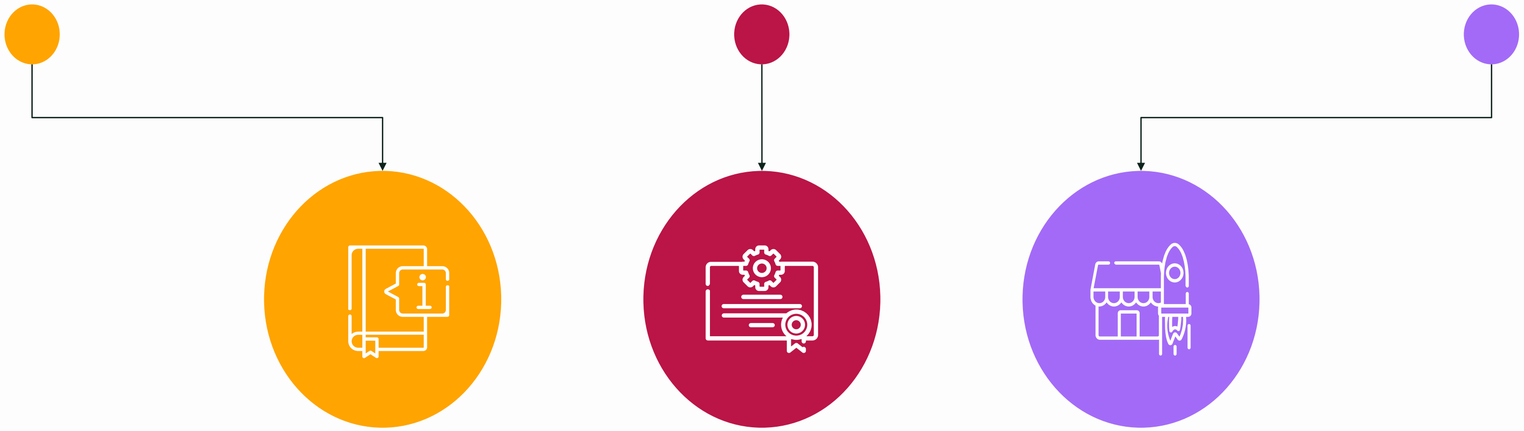
Punitive or aspirational ethical codes that set principles.

Regulation

Describes the conduct to be avoided and prescribes sanctions for those who do not take heed

Institutional Reform

Building ethics into the operations and decision making of the institution through its organizational design.



Technical vs. Organizational

Technical

Methods to instill human values in technical design.

Development of methodologies and assessments for the integration and auditing of values in design combined with efforts to provide normative frameworks for data scientists.

Organizational

Methods to embed ethics in organizational design: designing organizations capable of identifying, managing and mitigating ethical risks.

Development of practices on how to implement ethics in structures and processes that align with existing procedures and interests. Governance and ethical decision making of residual risks.

De Volksbank

Be in control of the ethical aspects of data science applications developed, deployed and used within De Volksbank.

New governance structure:

1. **Ethics Office** documents ethical aspects and decision-making for AI applications
2. **Ethical Advisory Board:** provides hand-on evaluation and approves the Ethical Review
3. **Ethics Committee** assists with normative judgements on specific ethical issues.

Alignment:

- Ethics office is aligned with Model Governance (focused on model risk) and Privacy Office (data use)

Ethical
Data
Science
Association

Why EDSA?

'What' to 'How'?

Still a large gap between principles and practice in AI ethics.

Organizational

Little actionable research that relates AI ethics to organizational praxis.

Holistic Approach

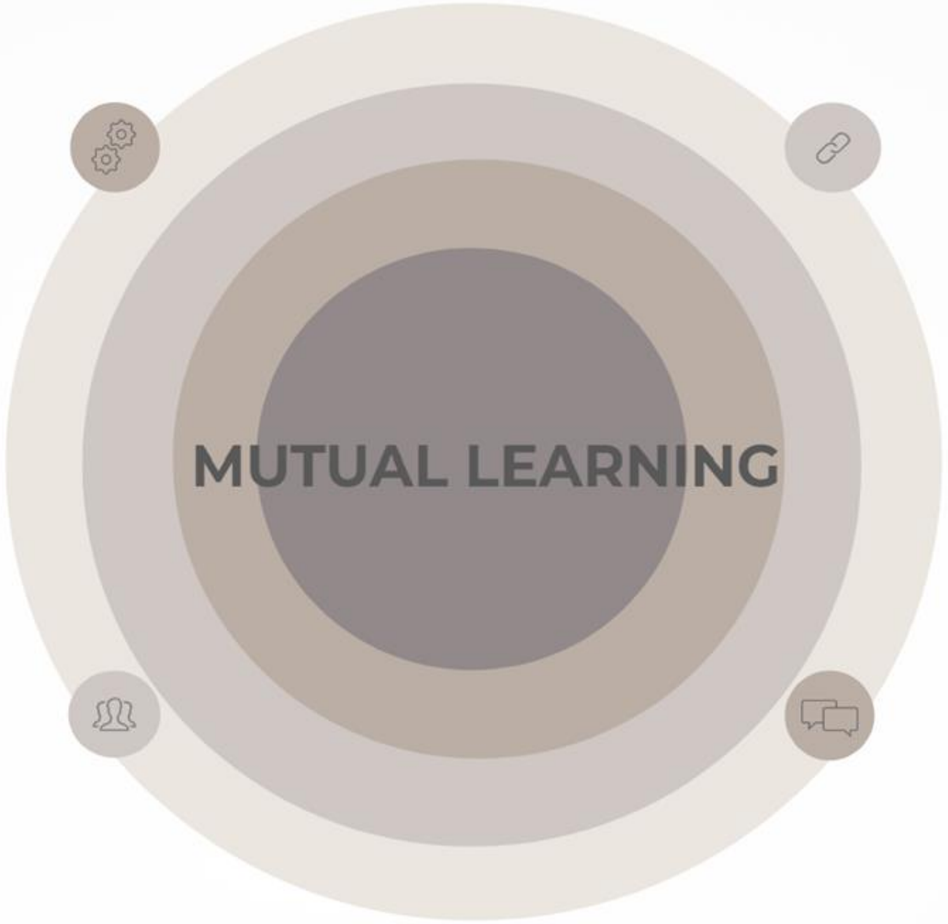
Ethical principles cannot be realized separately or sequentially; AI Ethics in organizations is often multi-dimensional.

Use context

The context of a use case is central in the discussion given the contextual nature of AI ethics.

Small-scale

Exploring themes and issues in small scale sessions so that everyone's expertise and experience can contribute.



Stakeholders

Bring together various groups of stakeholders (researchers, users, intermediaries, policy makers) from different sectors to learn about AI ethics challenges and solutions.

Interactive Learning

Facilitate an interactive learning process through mutual exposure of views and experiences, expectations and concerns.

Ethical Data Science Association

- **AI Impact**

AI Ethics Maturity

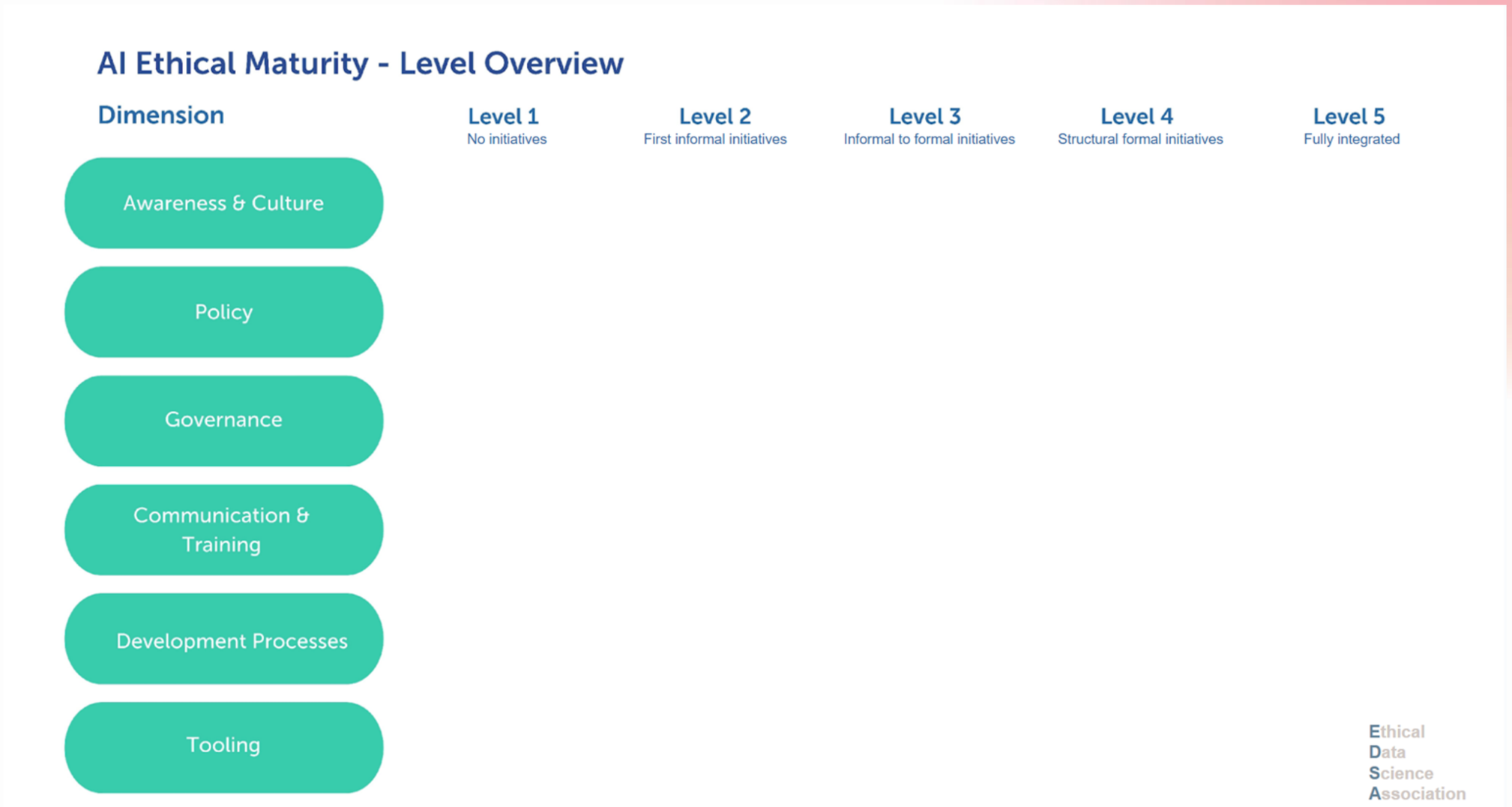
Six organizational dimensions

Ethical Data Science Association

sification: Internal

arisa

- **AI Impact**


AI Ethical Maturity - Level Overview
Dimension
Level 1
No initiatives
Level 2
First informal initiatives
Level 3
Informal to formal initiatives
Level 4
Structural formal initiatives
Level 5
Fully integrated
Awareness & Culture
Policy
Governance
Communication & Training
Development Processes
Tooling
Ethical Data Science Association

- **AI Impact**

Develop a roadmap of activities and steps

Dimension	Current & Target Maturity
Awareness & Culture	Current 2023 2026 2028 — Level 1 2 3 4 5
Policy	Current 2023 2026 2028 — Level 1 2 3 4 5
Governance	Current 2023 2026 2028 — Level 1 2 3 4 5
Communication & Training	Current 2023 2026 2028 — Level 1 2 3 4 5
Development Processes	Current 2023 2026 2028 — Level 1 2 3 4 5
Tooling	Current 2023 2026 2028 — Level 1 2 3 4 5

Steps To Take

Ethical Data Science Association

arisa

AI Impact

Awareness & Culture

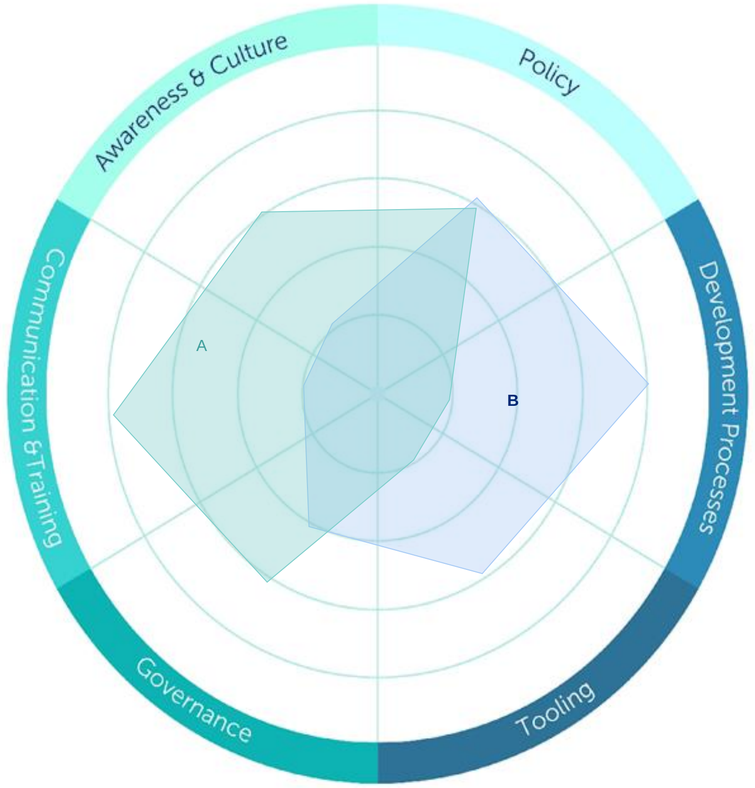
Policy

Communication & Training

Development Processes

Governance

Tooling



arisa

#AIskills4all | @AIskillsEU | linkedin.com/AIskillsEU

arisa

Thanks!

Feel free to reach out to:
LinkedIn: Joris Krijger
Email: krijger@esphil.eur.nl

Co-funded by the European Union

www.aiskills.eu

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Charting the Future – Exploring compliance readiness, safety innovation and open legal questions

Cornelia Kutterer, Managing Director, Considerati

Co-funded by the European Union

www.aiskills.eu

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

- About CONSIDERATI

Legal and public affairs consultancy

Legal services: Data & AI compliance

- Privacy, security and data-related ccompliance (DPOaaS, DPIAs, operational support for privacy compliance operations)
- AI Governance services (Risk assessment, HRIAs) in collaboration with MLOps tooling
- Training

Hybrid
Legal research
Regulatory dialogue / cooperation
Codes of conduct

www.considerati.com

Public Affairs

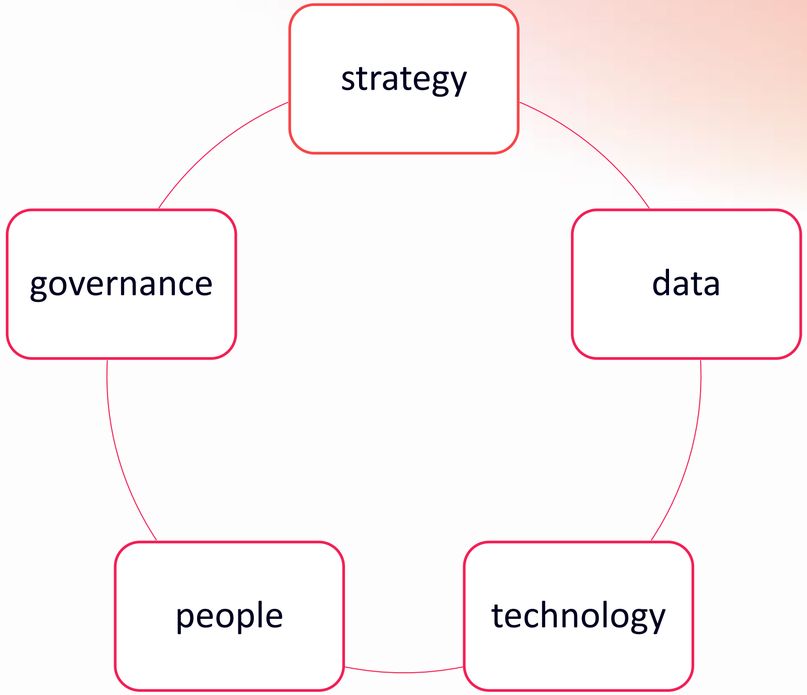
- Tech policy
- Healthcare
- Fintech

- **Overview**
- From Responsible AI Governance Frameworks to Compliance with the AIA
- Navigating the AI Act
- Tools (& procedures): MLOps for compliance
- DPIAs, FRIAs, due diligence & corporate responsibility

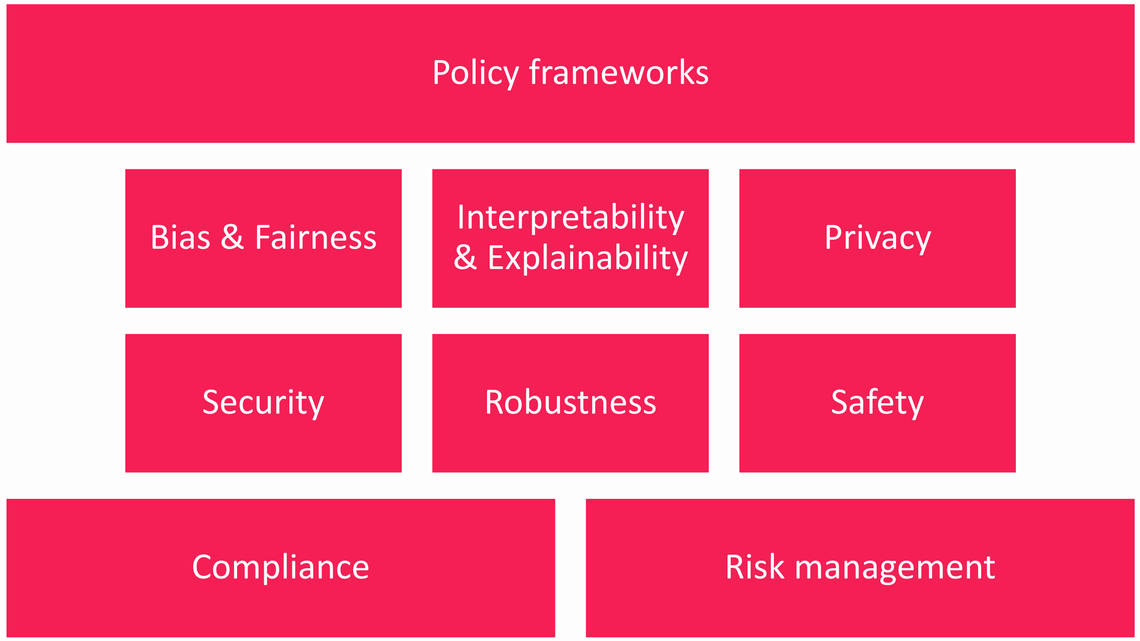
RAI as part of corporate responsibility

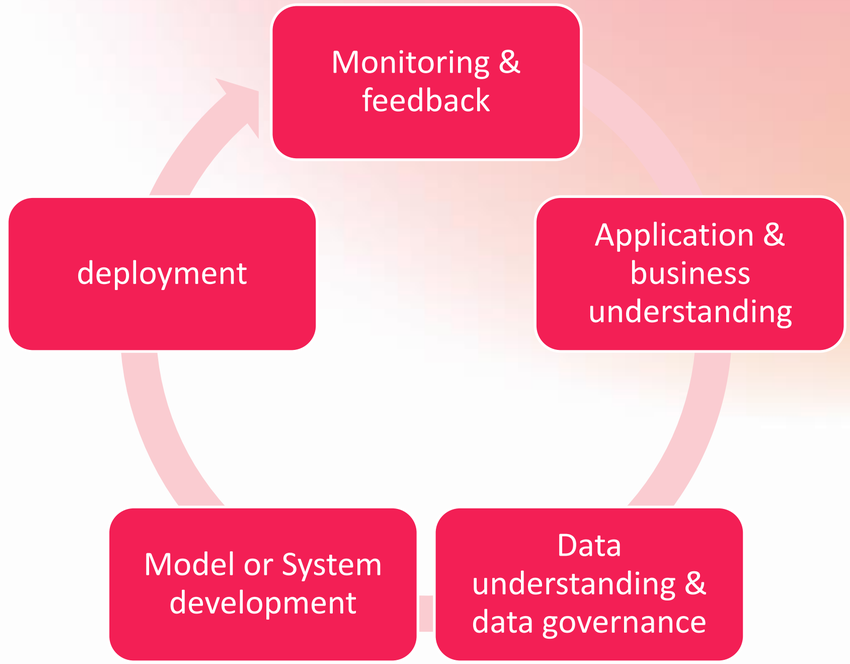
- **From Responsible AI Governance to Compliance with the AIA**
- What business leaders are concerned about
  - Risks to reputation with customers, compliance with growing list of regulations, making incorrect decisions based on AI, lengthy risk management processes, concerns re AI security vulnerabilities, inconsistent approaches managing AI risks across organizations
  - What is the AI maturity of an organization (degree to which organizations have mastered AI-related capabilities in the right combination to achieve high performance for customers, shareholders and employees)?
  - How do companies conceptualize the threats that their use of AI pose for individuals, groups and the broader society?
  - What substantive benchmarks, management processes and technological solutions do they use towards this end?

- **From Responsible AI Governance to Compliance with the AIA**
- AI Maturity
  - Responsible AI Maturity Model (RAI MM) are frameworks to help organizations identify their current and desired levels of RAI maturity.
  - Many RAI MM developed by leading companies



- **From Responsible AI Governance to Compliance with the AIA**
- Responsible AI Governance frameworks, tools and processes

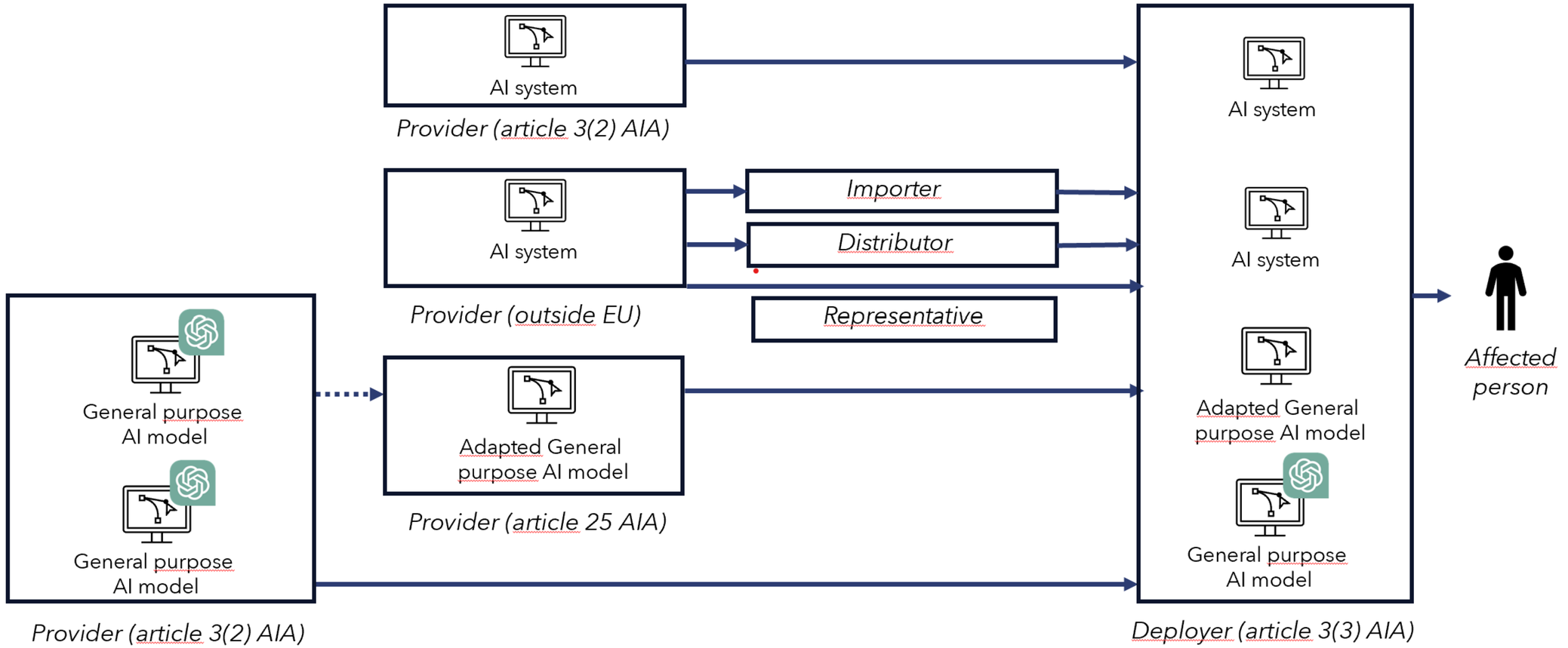




The dawn of a **new practice:** cross-disciplinary approach, starting with leadership

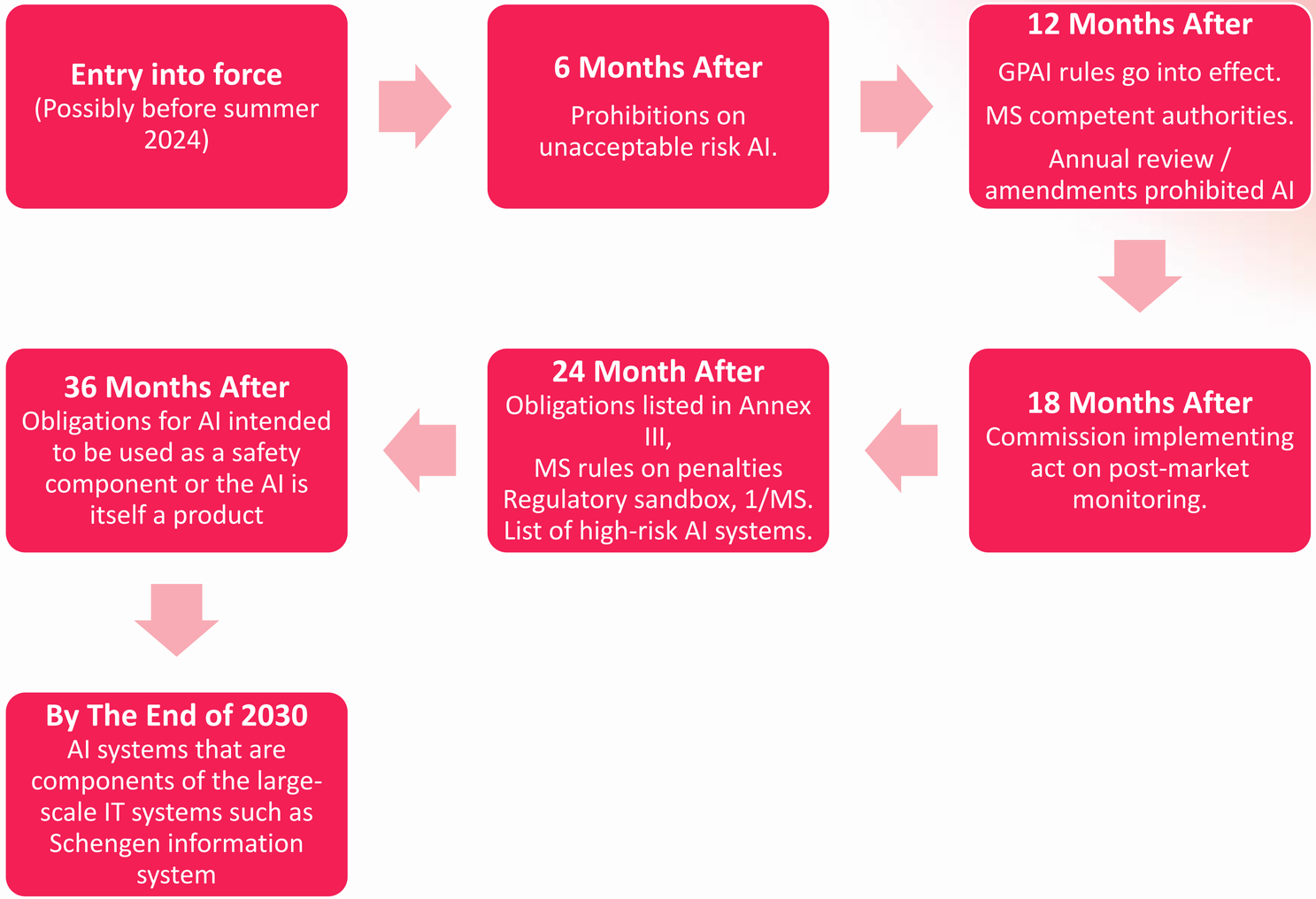
Navigating the AI Act

- Overview of regulated actors

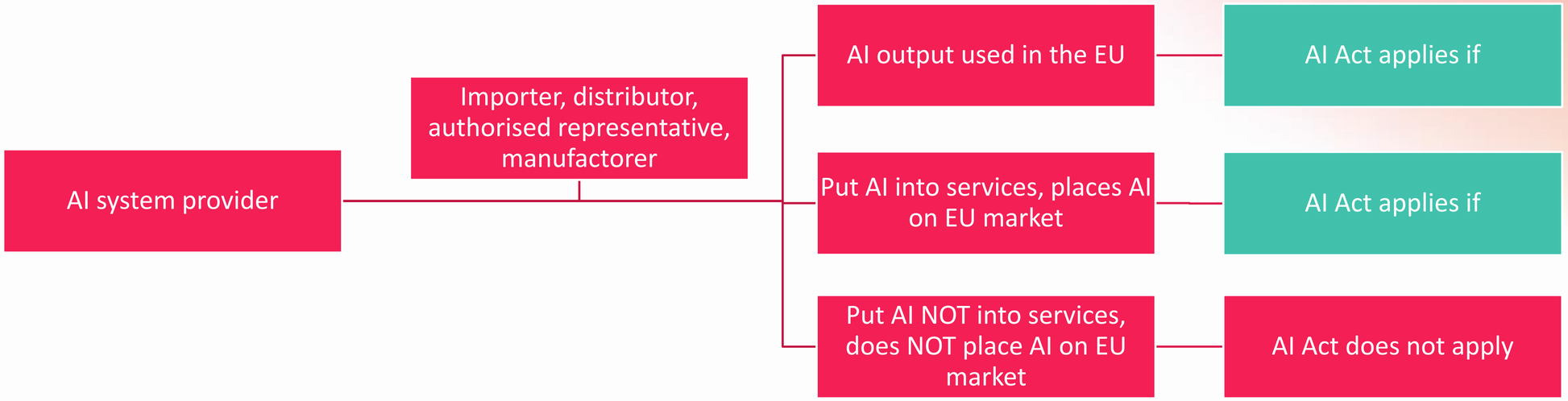


Navigating the AI Act

- timeline

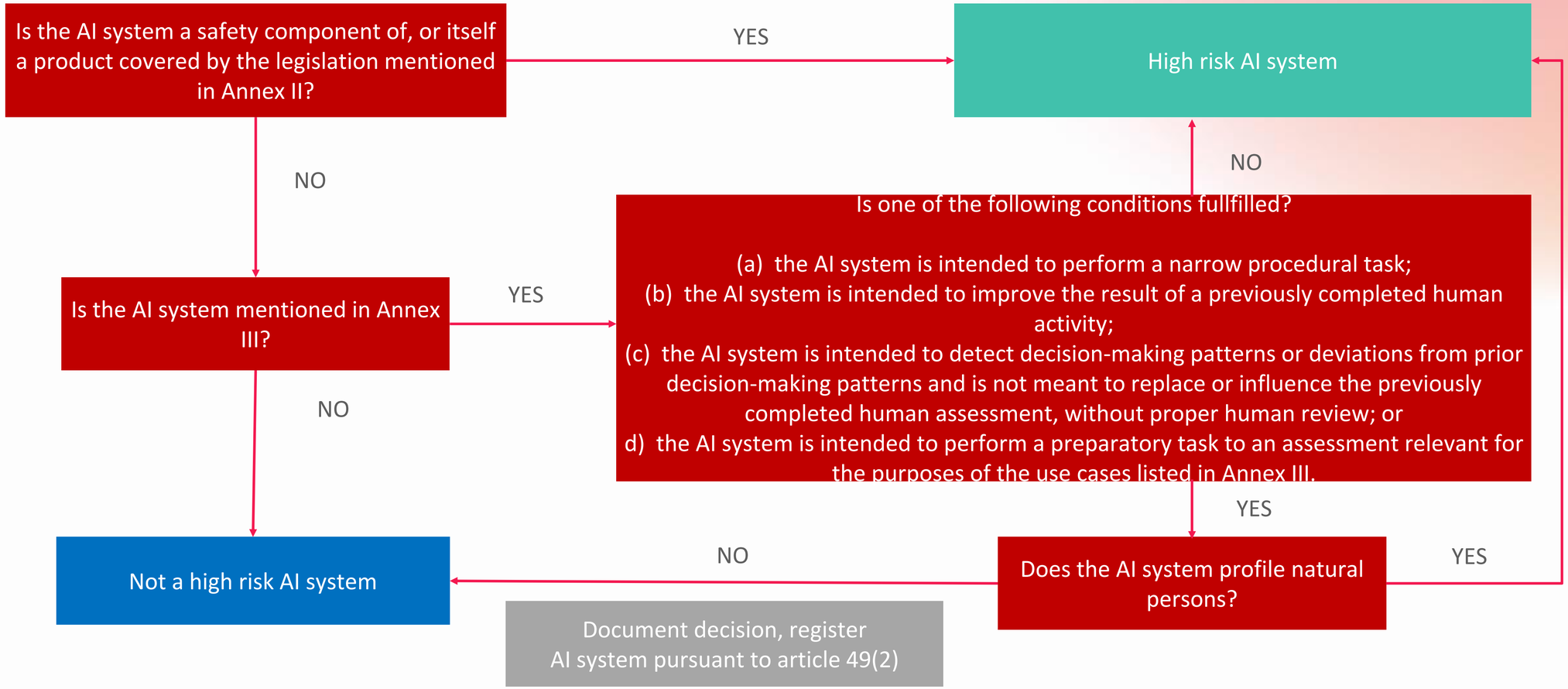


- **Navigating the AI Act**
  - AI System provider



Navigating the AI Act

High-risk AI system



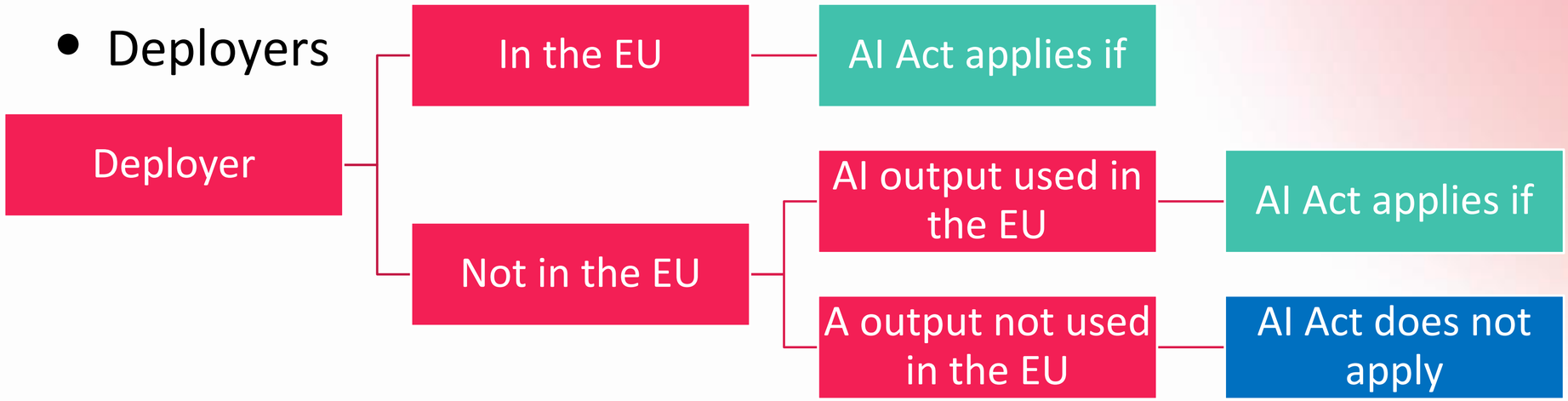
Navigating the AI Act

Obligations of AI system providers

Article 9 to 15	Article 16	Article 17
• Risk management system • Data and data givernance • Technical documentation • Record keeping (logs) • Transparency and instructions for deployers • Human oversight • Accuracy, robustness, cybersecurity	• Indicate name, contact address • Quality management • Conformity assessment • Conformity declaration • Affix CE marking • Registration • Corrective action / duty of information • Regulatory cooperation • Accessibility requirements	• Strategy for regulatory compliance • Design, design control and verification measures • Examination, test and validation measures • Technical specification • Data management system and procedures • Risk management system • Post market monitoring system • Serious incident reporting procedures • Communication with national authorities • Resource management • Accountability framework staff responsibility

Navigating the AI Act

- Deployers



Obligations applicable to all deployers

- Appropriate technical & organizational measures to follow instructions for use, ensure competence, training etc of human oversight, input data representative for intended purpose, monitoring and reporting of serious risks and incidents, keep logs generated, DPIAs, regulatory cooperation

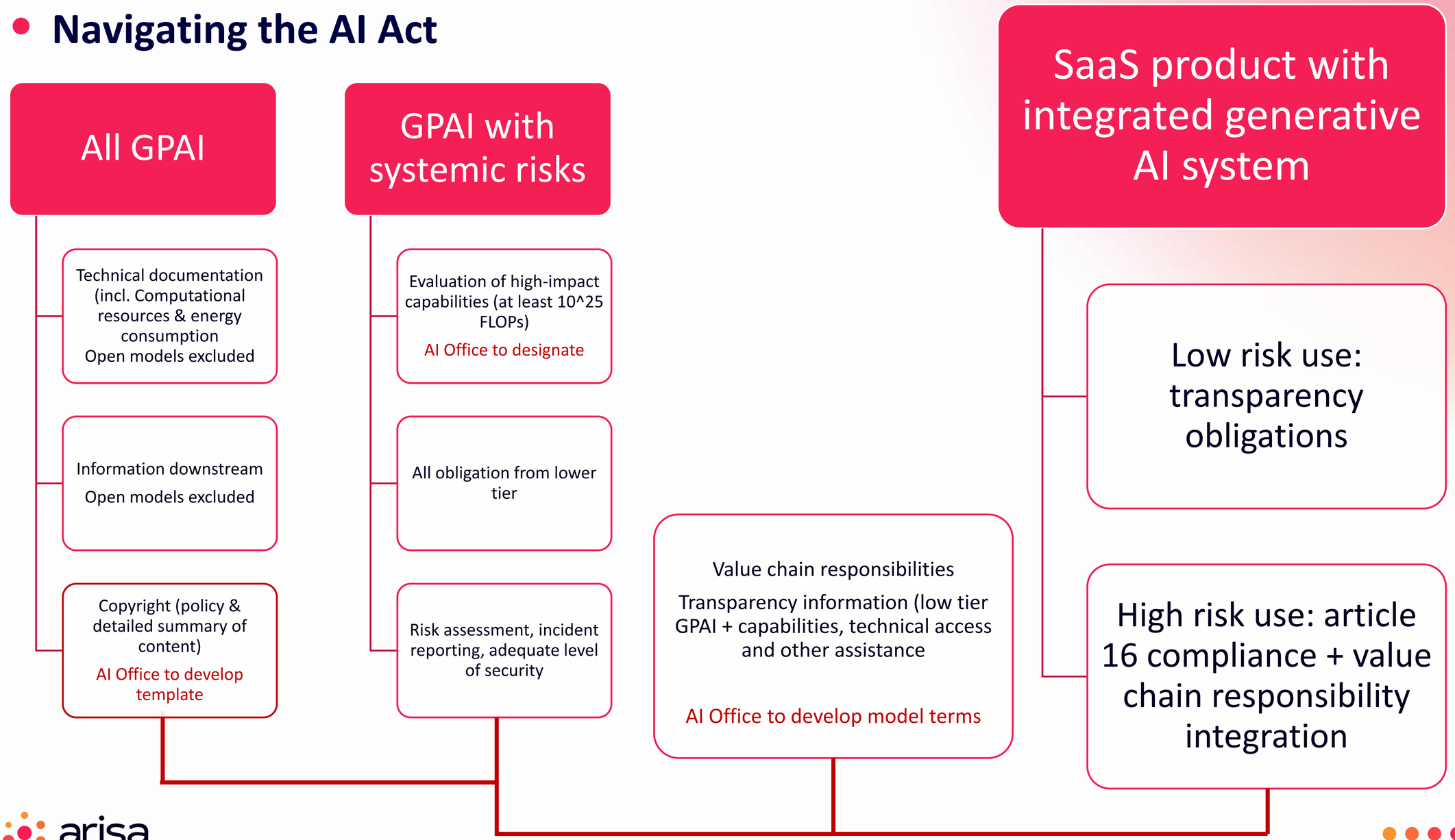
Additional obligations for specific criteria applies

- For workplace AI systems, inform workers' representative, if public authority register, if post-remote biometrics, seek authorization in criminal investigations, if public body, conduct Fundamental Rights Impact Assessment, If decision using Annes III, inform individual

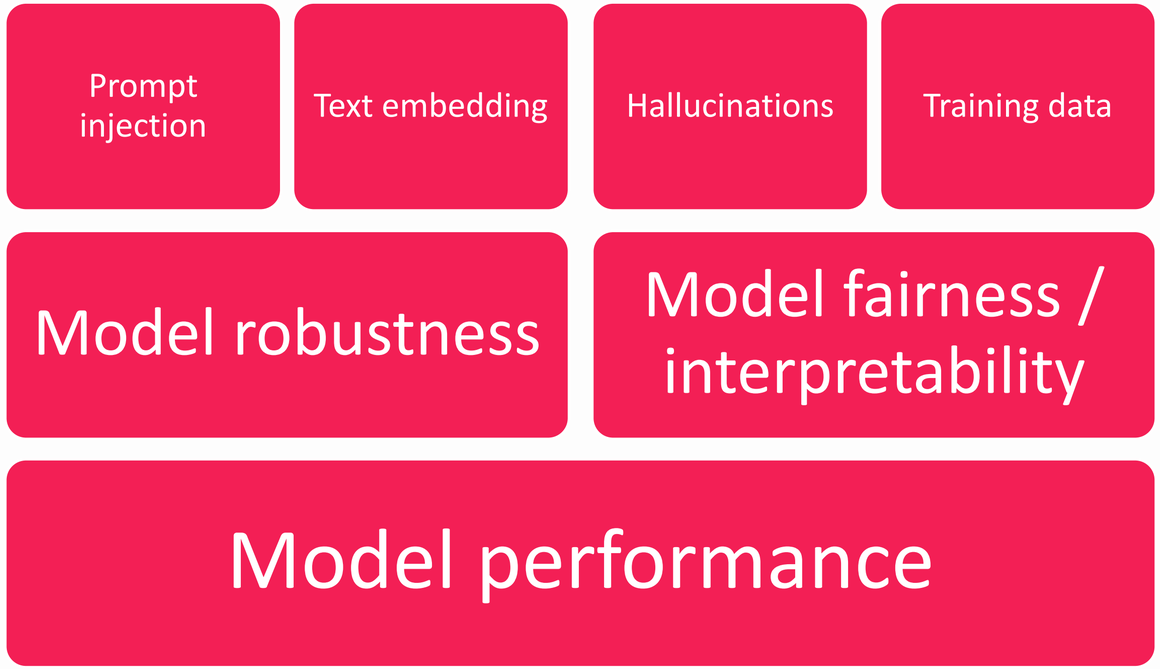
Fundamental Rights Impact Assessment

- Describe process in which AI system will be used, period of time/frequency, people/groups affected, specific harms, human oversight measures, risk mitigation measures, internal governance and complaint mechanism

Navigating the AI Act



- **Navigating the AI Act**
  - Challenges with GPAI integration



Related to state of the art

models that are resilient to adversaries, unusual situations, Black Swan events

Detection of malicious use, discover unexpected model functionality

Models that safely optimize for hard-to-spcify human values

Related to AI Act safeguards:

GPAI rules exclude narrow AI

Models that are not designated as highly capable not under the same obligations towards value chain despite potential use in sensitive or high-risk use cases

Other legal concerns

Copyright infringement

Data leaks

Disinformation

Tools (& procedures) MLOps for compliance

1. **Conduct a comprehensive AI audit :** Assess your current AI systems and processes to determine how they align with the EU AI Act. Identify areas that require changes or enhancements to meet compliance standards.

2. **Develop a risk management strategy:** For high-risk AI applications, establish a robust risk management framework. Include mechanisms for monitoring, reporting, and mitigating risks associated with AI systems. For low-risk applications, maintaining transparency in their operations and ensuring the accuracy of the information it processes is still crucial.
   Clearly communicate on how the AI functions and the nature of data it handles.

3. **Invest in training and awareness:** Ensure your staff is well-informed about the EU AI Act and its implications. Regular training sessions can help build a compliance-focused culture within the organisation.

4. **Engage with AI ethics and compliance experts:** Consult AI ethics and compliance experts to navigate the complex regulatory environment effectively. They can provide insights into best practices and help you stay ahead of regulatory changes.

5. **Foster transparency and accountability:** Develop clear policies and procedures for AI transparency and accountability. Maintain detailed records of AI decision-making processes and outcomes.

6. **Leverage technology for compliance:** Use general and AI-specific compliance management software and tools to streamline and automate parts of your compliance processes, making them more efficient and less prone to errors.

7. **Stay informed and agile:** Keep abreast of regulatory updates and be prepared to adapt your AI strategies as the regulatory landscape evolves.

- **Tools (& standards) MLOps for compliance**

- Platforms have developed MLOps tools to help with implementing and scaling responsible AI governance (some open-sourced)

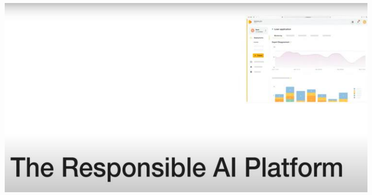
- Many companies that provide these services emerge – still nascent market

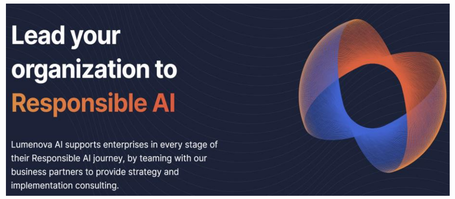
- Deeploy, BreezeML, Lumenova, and many others

- Reponsum, OneTrust

ISO standards 42001

Cen Cenelec JTC121





arisa

#AIskills4all | @AIskillsEU | linkedin.com/AIskillsEU

Co-funded by
the European Union

www.aiskills.eu

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Let’s stay in touch

- Visit the ARISA website for a full list of the project’s resources, blog articles, interviews with the Alliance members & more: https://aiskills.eu/

- Follow us on social media for regular updates on the project:
  LinkedIn: @ARISA – AI Skills
  X: @AIskillsEU

- Subscribe to the ARISA bi-monthly newsletter to get the latest news, releases & insights from the project delivered to your inbox.

arisa

#AIskills4all | @AIskillsEU | linkedin.com/AIskillsEU

Co-funded by
the European Union

www.aiskills.eu

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.