

# Recommender systems research under the EU's Digital Services Act and AI Act

*João Vinagre, Emilia Gómez, David Fernández Llorca*

FAccTRec 2024

RecSys 2024, Bari, October 14<sup>th</sup>

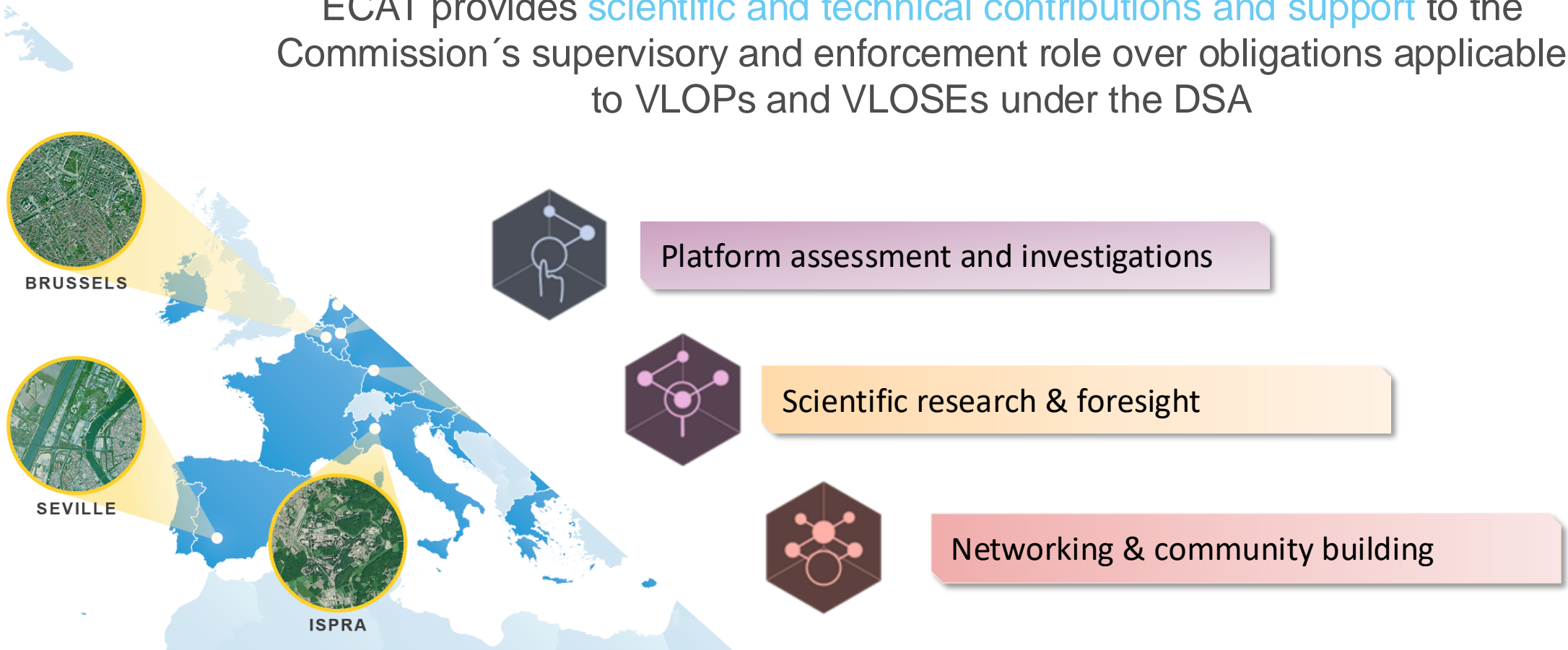
# About the JRC

The JRC provides independent, evidence-based science and knowledge, supporting EU policies to positively impact society.



# About ECAT

ECAT provides **scientific and technical contributions and support** to the Commission's supervisory and enforcement role over obligations applicable to VLOPs and VLOSEs under the DSA



# HUMAIN: AI and human behaviour



# Knowledge synthesis



Computers in Human Behavior Reports

Volume 14, May 2024, 100404



Review

## Social Media Use and adolescents' mental health and well-being: An umbrella review

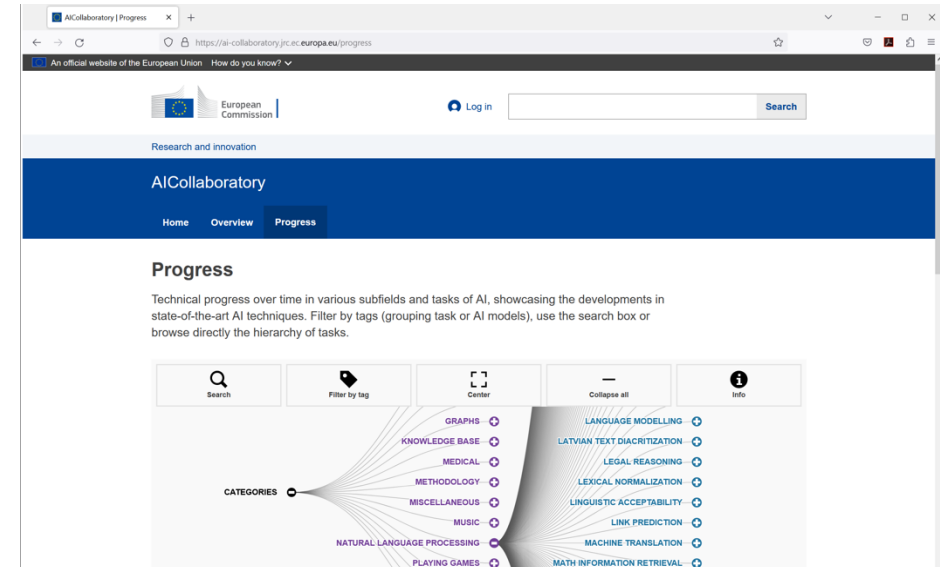
Arianna Sala <sup>a</sup>  , Lorenzo Porcaro <sup>b</sup>  , Emilia Gómez <sup>a</sup>  

- **Criteria:** Population 10-19 y-o; Exposure Social Media; Outcomes: Well-being; Il-being; Mental health;
- **Corpus:** 1470 publications -> 24 high quality articles (~80% >2020);
- **Intervening factors:** I) Individual psycho-social-factors; ii) Individual use of SM; iii) Platform functioning;
- **Risk mitigation:** Sharing responsibilities from individuals to platforms under the [Digital Services Act](#).



# AI mapping and indicators

- EU Coordinated plan on AI
- Artificial Intelligence Indicators
  - Progress, benchmarks
  - Country
  - Academia vs industry
  - Gender
- EU AI Watch Index
- AI Office



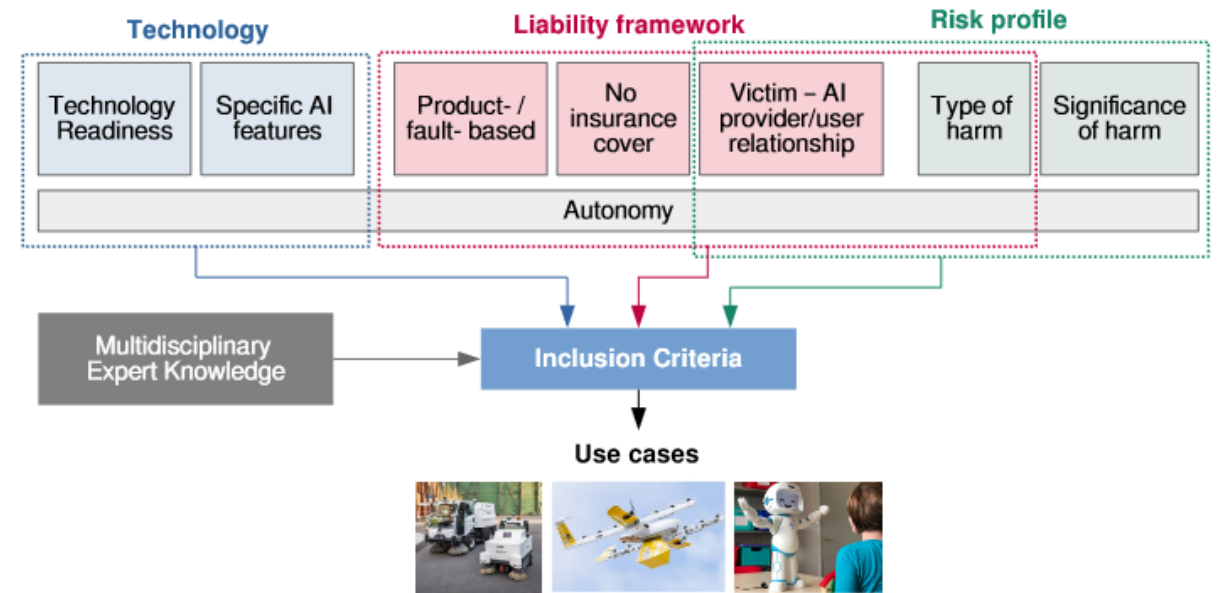
EC Diversity and  
Inclusion Awards 2023

[https://ai-watch.ec.europa.eu/humaint/divinai\\_en](https://ai-watch.ec.europa.eu/humaint/divinai_en)

# System-centric evaluation/auditing



- AI liability directive
- Analysis of use cases of the burden of proof



## Liability Regimes in the Age of AI: : a Use-Case Driven Analysis of the Burden of Proof

Authors: [David Fernández Llorca](#), [Vicky Charisi](#), [Ronan Hamon](#), [Ignacio Sánchez](#), [Emilia Gómez](#) | [Authors Info & Claims](#)

Journal of Artificial Intelligence Research, Volume 76 • <https://doi.org/10.1613/jair.1.14565>

Published: 04 May 2023 [Publication History](#)



# Human-centric evaluation



2019  
N=20  
Impact on children's learning



2020-2021  
N=84  
Impact on group's social behaviour  
Impact on children's perception and trust

Guidelines for developers

2023  
N = 65 (49 children, 16 Adults)  
Transparency, guardian consideration and age adapted vocabulary improve trustworthiness of conversational agents



Charisi, V., Gomez, E., Mier, G., Merino, L., & Gomez, R. (2020). Child-Robot Collaborative Problem-Solving and the Importance of Child's Voluntary Interaction: A Developmental Perspective. *Frontiers in Robotics and AI*, 7, 15.

Charisi V., Merino, L., Caballero, F., Escobar, M., Gomez, R., Gomez, E. The Effects of Robot Cognitive Reliability and Social Positioning on Child-Robot Team Dynamics. International Conference on Robotics and Automation (ICRA2021).

Escobar, M., Charisi, V., Gómez, E. I've seen a robot!: The impact of cognitive reliability and expressivity in children's perception of a robot. ACM on Human-Computer Interaction, 6(CSCW2), 1-23.





# DSA Delegated Act on Data Access

Digital Services Act. **Technical conditions** under which VLOPs and VLOSEs are to share data with vetted researchers.

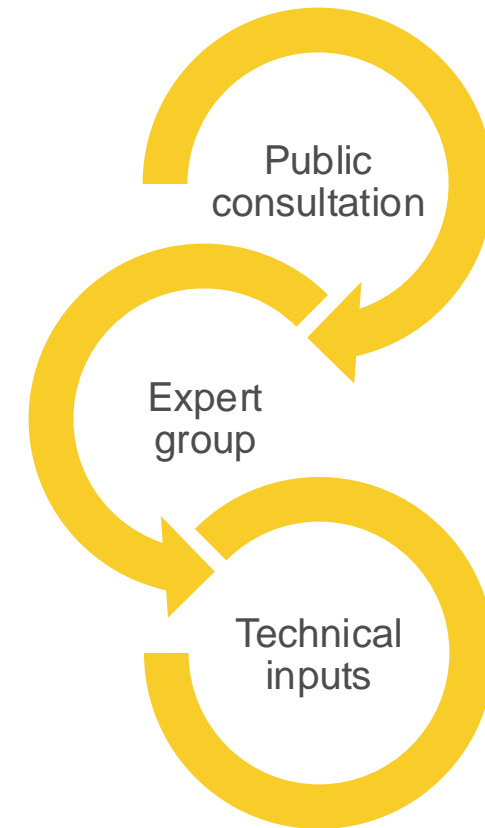
## 1. Provide technical inputs to the DA

- Study the current data access landscape
- Identify potential gaps and bottlenecks in implementation
- Provide adequate technical means to overcome potential issues

## 2. Community building

- Mapping of researchers/groups, projects (e.g. ERC grants)
- Dissemination (workshops, info sessions)
- Promote the involvement of researchers of identified communities

[https://algorithmic-transparency.ec.europa.eu/news/faqs-dsa-data-access-researchers-2023-12-13\\_en](https://algorithmic-transparency.ec.europa.eu/news/faqs-dsa-data-access-researchers-2023-12-13_en)

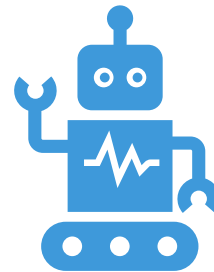


## 2 AI Regulations in the EU



### **Digital Services Act**

In force since November 2022



### **AI Act**

In force since August 2024

# Ethics guidelines for trustworthy AI - 2019

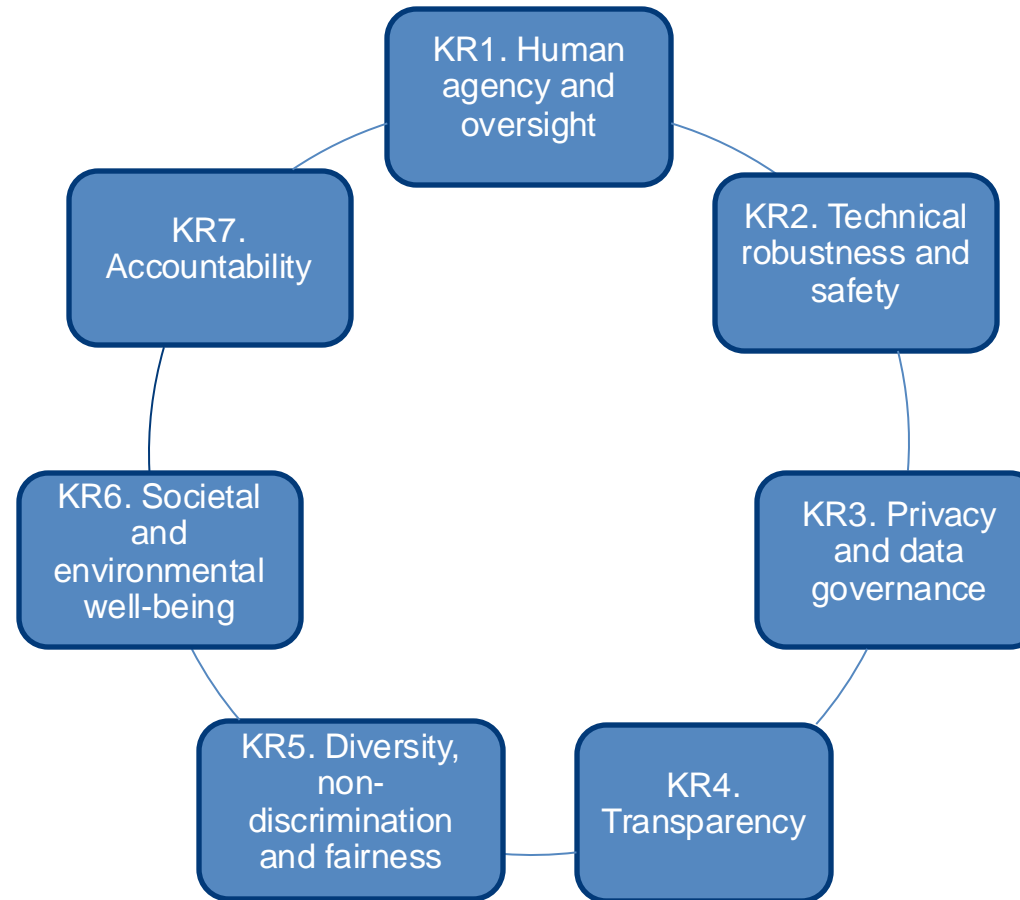
**Lawful** - respecting all applicable laws and regulations.

**Ethical** - respecting ethical principles and values.

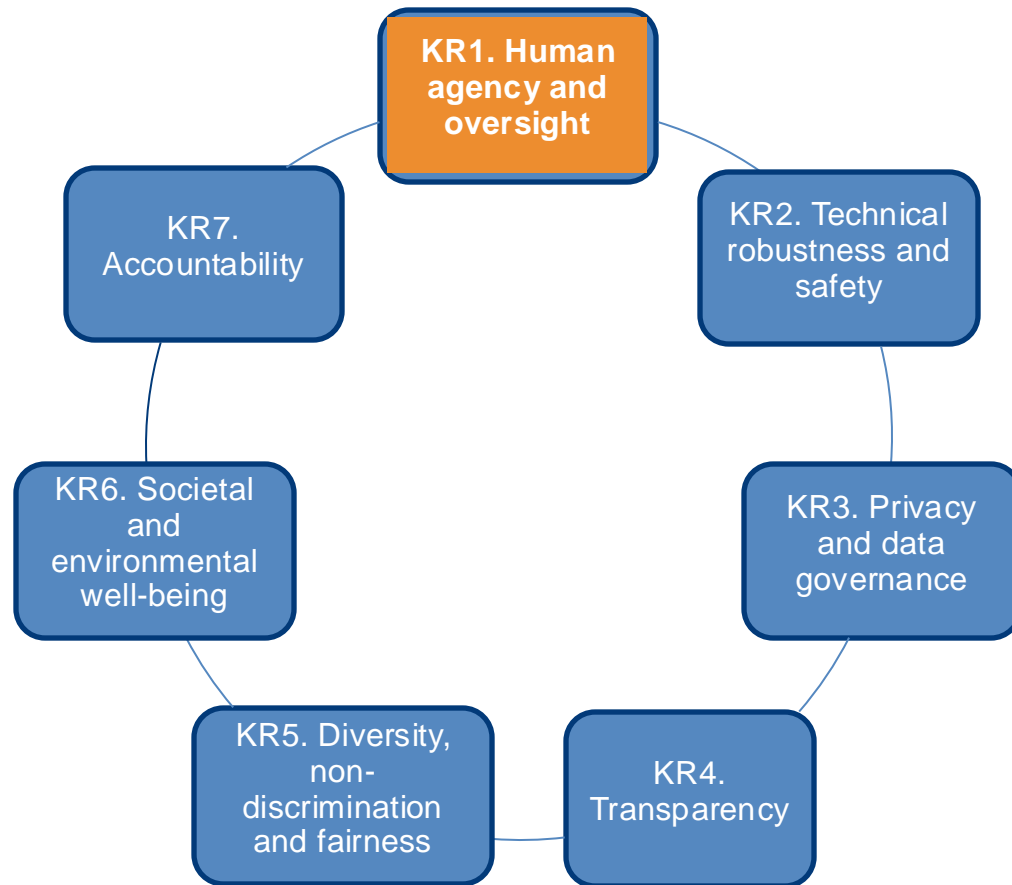
**Robust** - both from a technical perspective while taking into account its social environment. Avoid intentional/unintentional harm.



# Key requirements for trustworthy AI

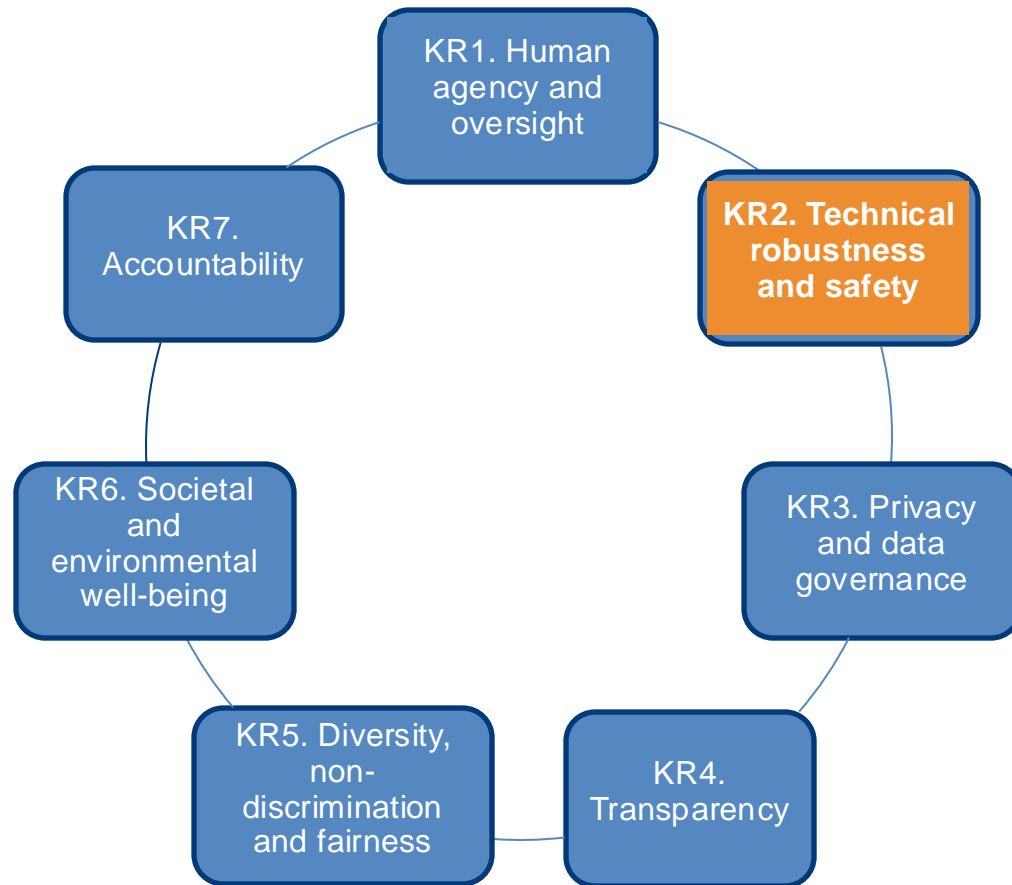


# Key requirements for trustworthy AI



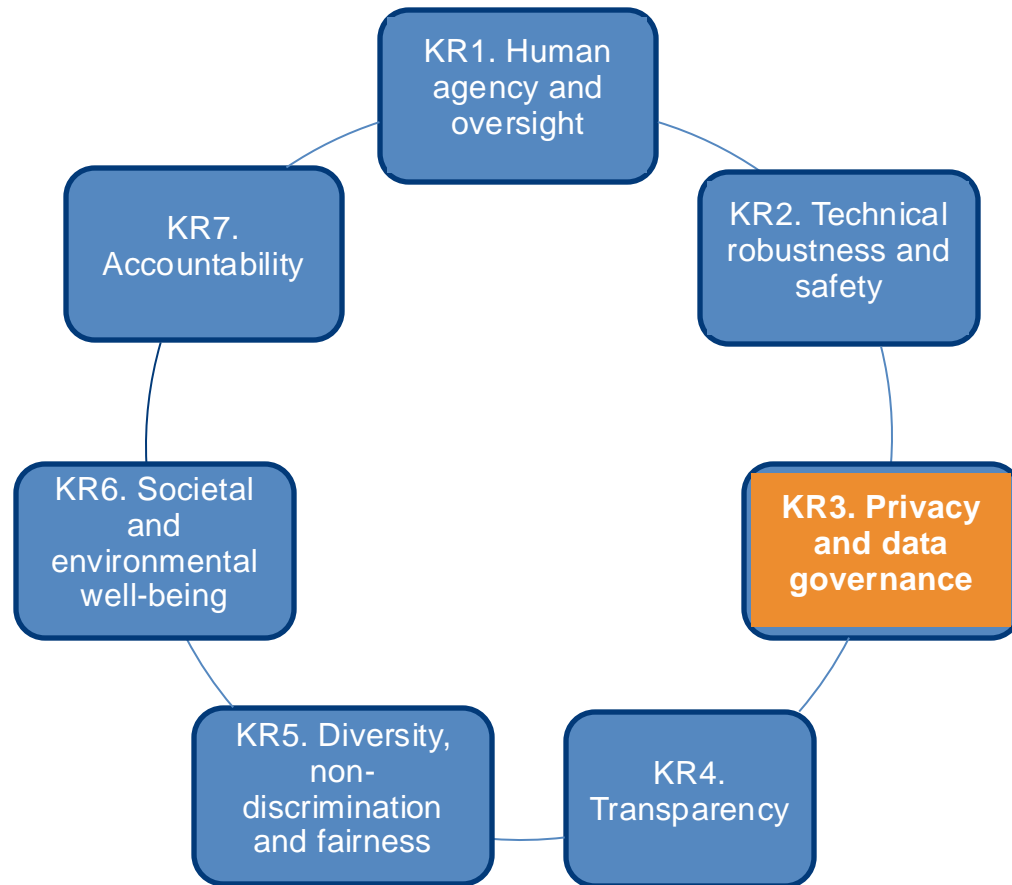
- Respect for **fundamental rights**
- Human agency (autonomy)
- Human oversight

# Key requirements for trustworthy AI



- Resilience to attack
- Fallback plan
- Accuracy guarantees
- Reliability
- Output reproducibility

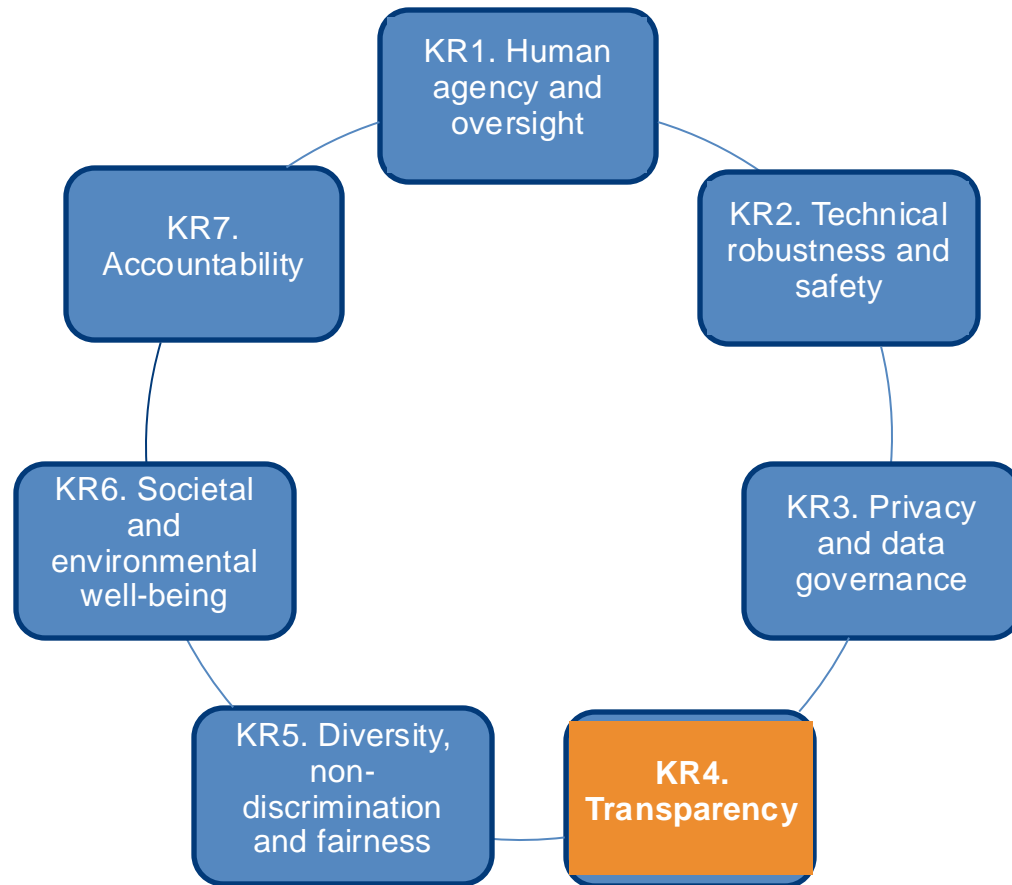
# Key requirements for trustworthy AI



- Privacy and data protection
- Data quality and integrity
- Data access protocols

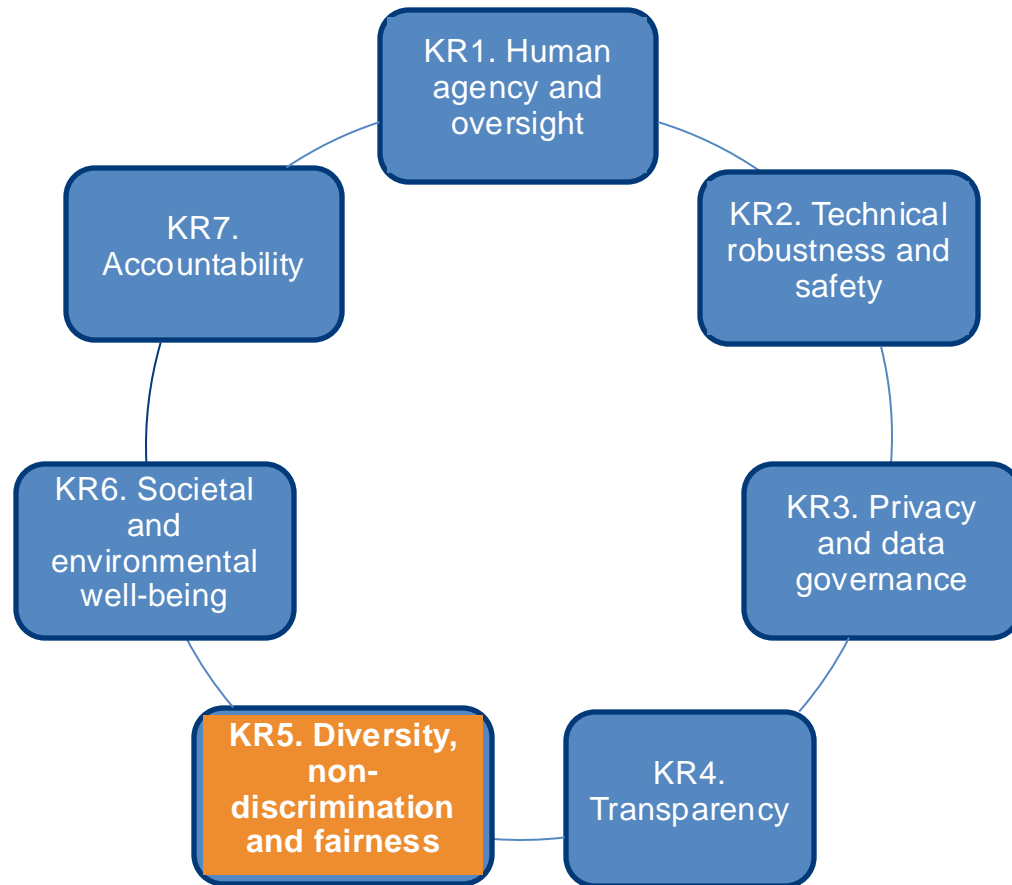


# Key requirements for trustworthy AI



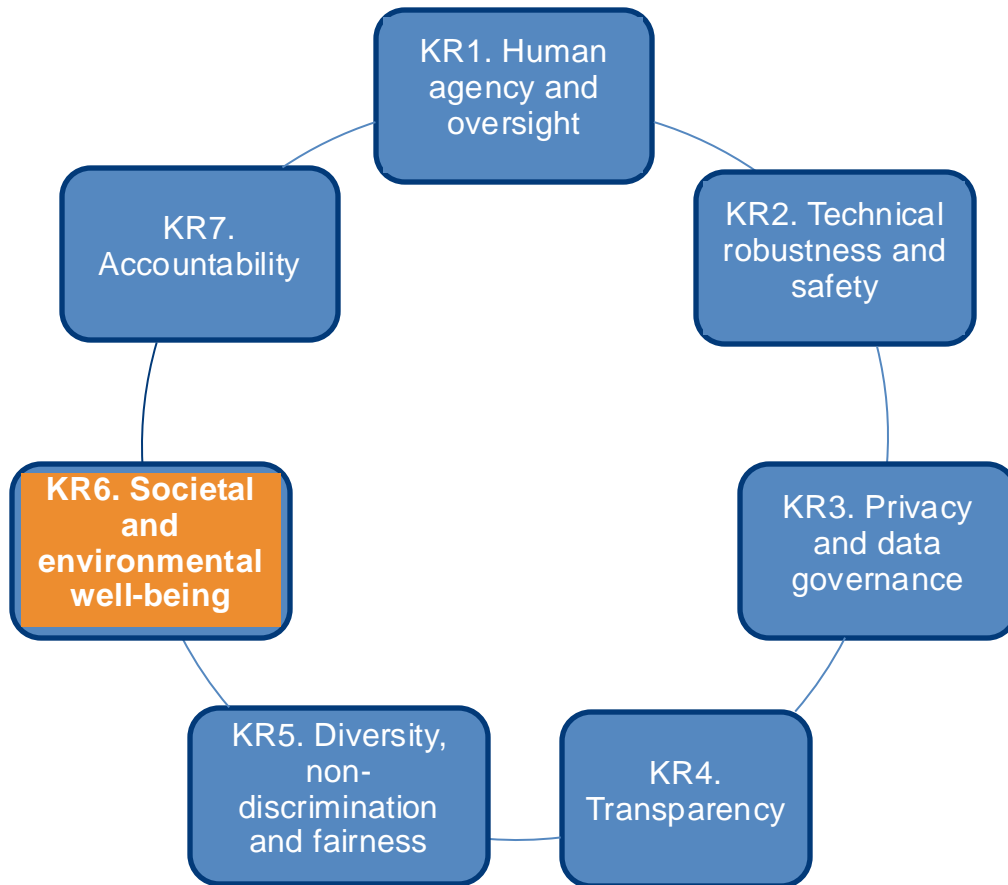
- Traceability
- Explainability
- Communication (identifiability and clarity of limitations)

# Key requirements for trustworthy AI



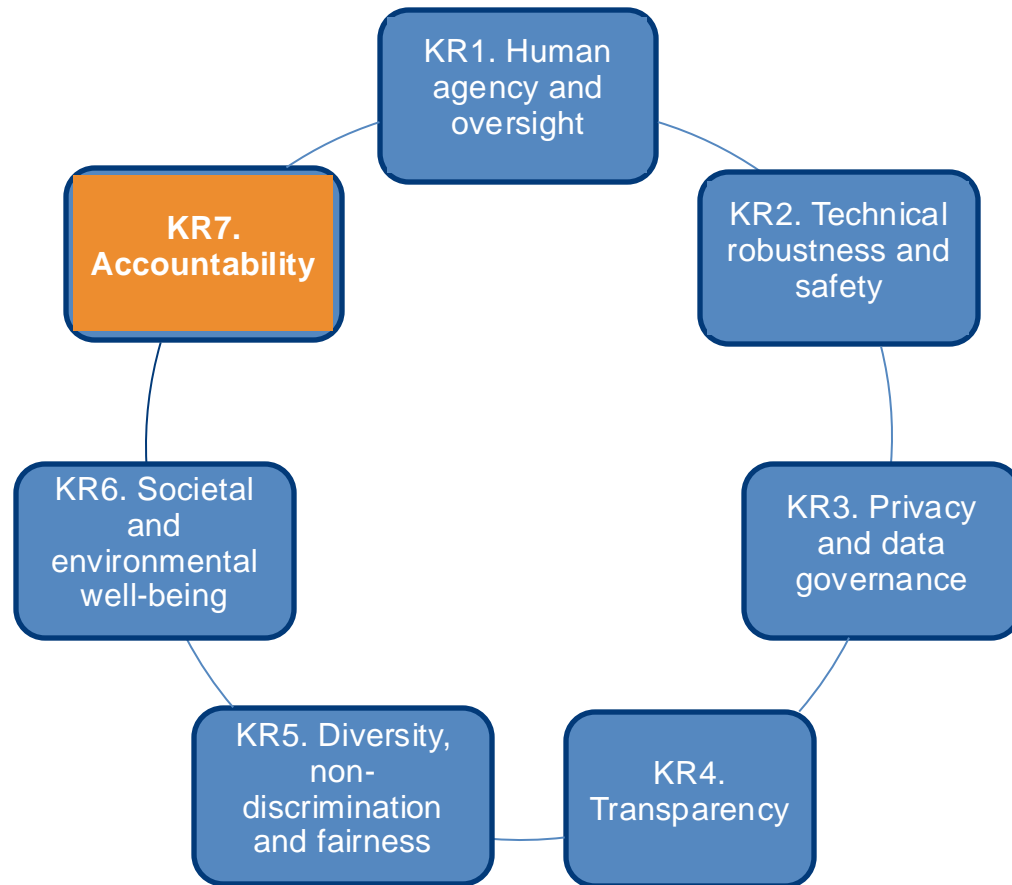
- Avoidance of unfair bias
- Accessibility and universal design
- Stakeholder participation

# Key requirements for trustworthy AI



- Environmental sustainability
- Social impact
- Society and democracy

# Key requirements for trustworthy AI



- Auditability
- Mitigation and reporting of negative impacts
- Trade-offs between requirements
- Redress

# Fundamental rights in the EU



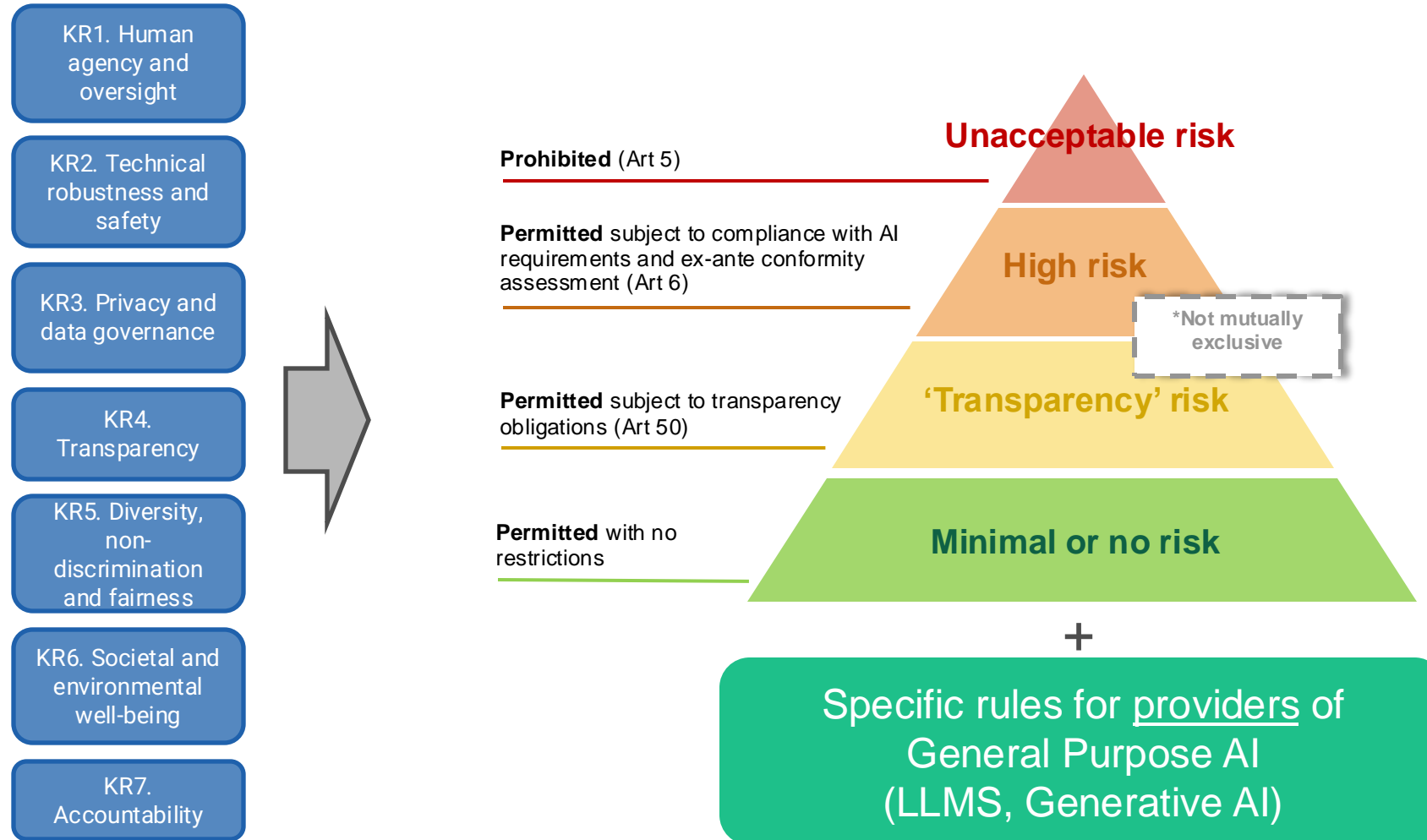
- Dignity
- Freedoms
- Equality
- Solidarity
- Citizen's rights
- Justice

**105** occurrences in the AI Act (144 pages)

**39** occurrences in the DSA (103 pages)

# AI Act

# AI Act: risk-based approach





# AI Act: prohibited practices (Art. 5)

Deceptive, manipulative, or exploitative AI, susceptible of causing harm

Social scoring

Predictive policing based on personality traits or characteristics

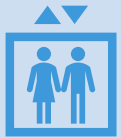
Facial recognition systems trained on indiscriminately scraped data

Emotion recognition at work or study places (except for health/safety)

Biometric categorisation (AI to deduce beliefs, sexual orientation,...)

Real-time remote biometric identification\*

# AI Act: High-risk systems (Art. 6)



AI is used in/as a safety component of a regulated product

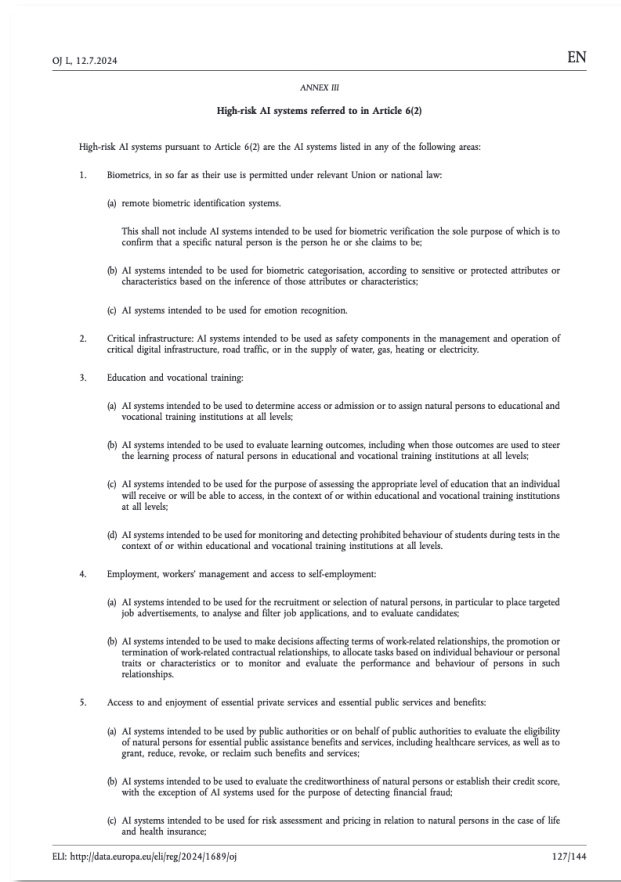


AI system is itself already a regulated product



AI application is listed in Annex III

# AI Act: the famous Annex III



- Biometrics (remote ID, categorisation, emotion recognition)
- Applications for critical infrastructures
- Education and training (admissions, evaluation, placement, monitoring in tests,...)
- Employment (selection, recruitment, task allocation, performance monitoring, promotions, lay-offs,...)
- Eligibility to essential services (public and private)
- Law enforcement (risk assessment, polygraphs, profiling,...)
- Migration, asylum and border control management
- Justice and democratic processes

# High-risk AI systems: requirements

- Risk management (pre- and post-market)
- Data governance
- Technical documentation (not necessarily public)
- Record-keeping: logs
- Transparency: documentation to deployers
- Human agency: mechanisms to monitor and take action while in production
- Accuracy, robustness and cybersecurity

# AI Act: General Purpose AI Models (GPAI)

## “General Purpose AI Model”

*means an AI model, including when trained with a large amount of data **using self-supervision at scale**, that **displays significant generality** and is **capable to competently perform a wide range of distinct tasks** regardless of the way the model is **placed on the market** and **that can be integrated into a variety of downstream systems or applications**.*

generality

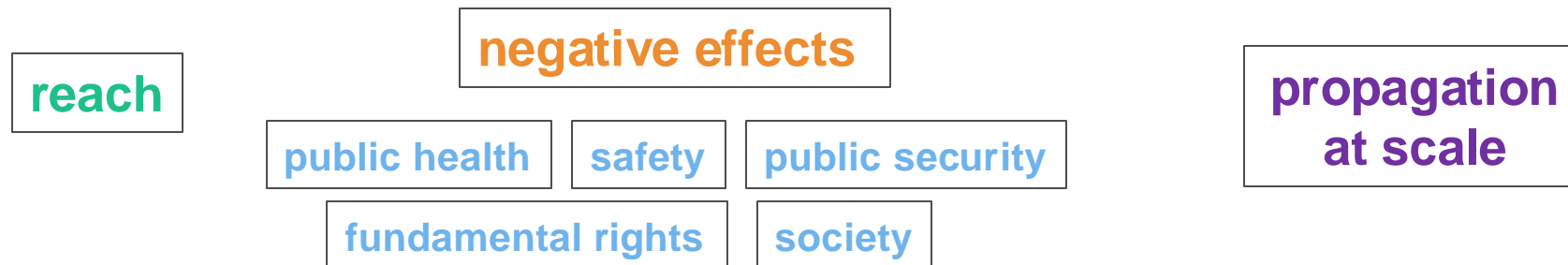
competently perform

variety of downstream  
systems and applications

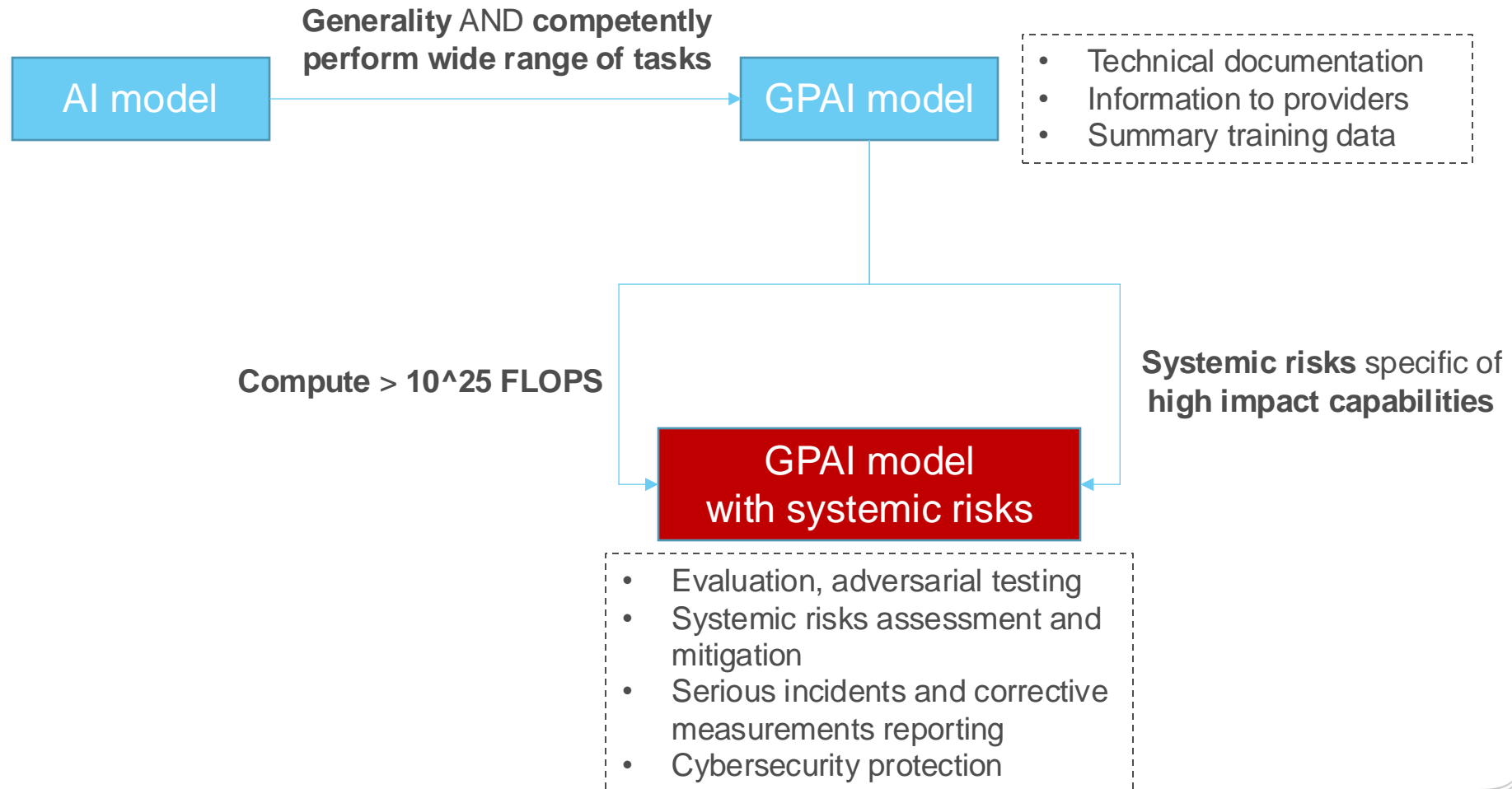
wide range of distinct tasks

# AI Act: Systemic risks

- ‘**systemic risk at Union level**’ means a risk that is **specific to the high-impact capabilities** of general-purpose AI models, having a **significant impact** on the **internal market due to its reach**, and with **actual or reasonably foreseeable negative effects** on **public health, safety, public security, fundamental rights, or the society as a whole**, that can be **propagated at scale** across the value chain;



# AI Act: GPAI models





# What should recsys providers worry about?

A few things:

- 🚫 Prohibited applications of recsys
- ⚠️ High-risk applications of recsys
- ⚠️ Using a self-developed GPAI model as (or together with) a recsys

*Recital (97): “(...) When the provider of a general-purpose AI model integrates an own model into its own AI system that is made available on the market or put into service, that model should be considered to be placed on the market and, therefore, the obligations in this Regulation should continue to apply in addition to those for AI systems (...)”*

# AI Act: Should researchers be worried?

## Article 2:

6. This Regulation **does not apply to AI systems or AI models**, including their output, specifically developed and put into service **for the sole purpose of scientific research and development**.

8. This Regulation **does not apply to any research, testing or development activity** regarding AI systems or AI models **prior to their being placed on the market** or put into service. Such activities shall be conducted in accordance with applicable Union law. **Testing in real world conditions shall not be covered by that exclusion.**

# AI Act: What opportunities for research are there?

- Risk assessment and mitigation
  - Evaluation, monitoring, metrics (offline and online)
  - Regulatory sandboxes, auditing
- Data governance
  - Privacy, security, interoperability, minimisation
- Collaboration with AI Office
  - Scientific panel, Advisory Forum

# Digital Services Act (DSA)

US &amp; WORLD / ENTERTAINMENT / SPOTIFY

## Spotify and Deezer ask EU regulators to stop Apple from abusing its dominance

The music Tech titans

By BEN POPPER  
Dec 13, 2017, 10:39 PM



# Google's algorithm shows prestigious job ads to men, but not to women

This isn't the first time that algorithm systems have appeared to be sexist

Julia Carpenter • Tuesday 07 July 2015 14:36 • Comments



in seinem Kampf gegen  
gestärkt. Das ist wichtig  
bekommen die Macht

An ex-YouTube insider  
recommendation algorithm  
divisive clips and conspir  
harm Hillary Clinton's b

Accueil | Les Marchés **LIVE** | Mon Argent

**LEcho**

An Van Gerven prend les commandes de Pfizer Belgique

Minoryx demande une autorisation de mise sur le marché pour le Ieriglitazone

Les utilisateurs puisent dans le

ABONNEZ-VOUS | LOGIN | JOURNAL | ? | Q

ACTU > ENTREPRISES > TECHNOLOGIE

TRIBUNE

Où en est-on en matière de régulation des plateformes numériques?

# DSA: pre-regulation

- Germany (NetzDG)
- France (Avia Law)
- Denmark
- Poland
- Hungary



# Proposal of DSA

KR1. Human  
agency and  
oversight

KR2. Technical  
robustness and  
safety

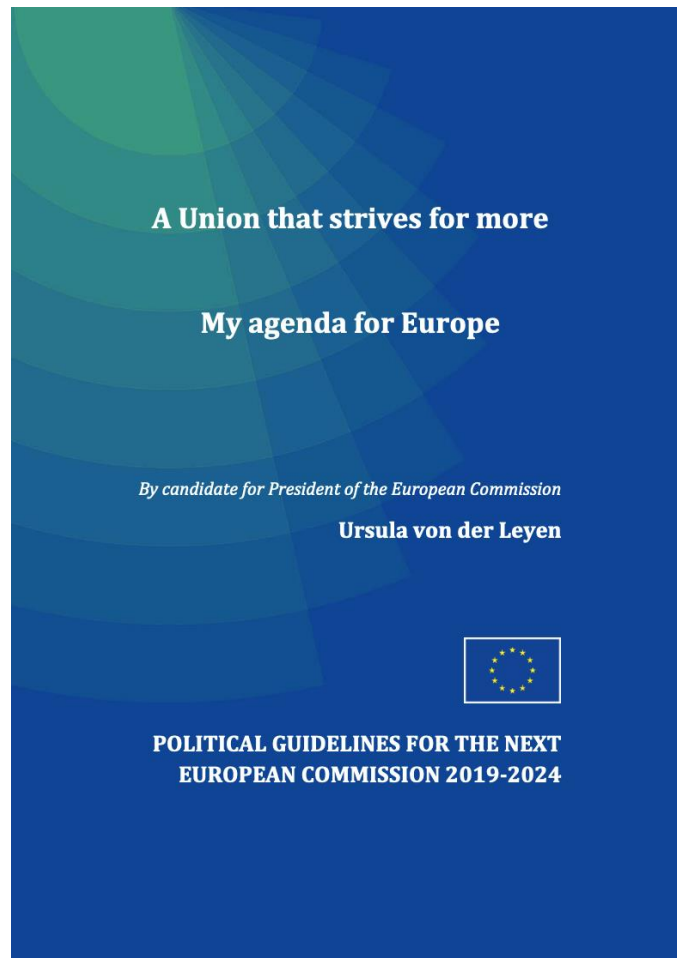
KR3. Privacy and  
data governance

KR4.  
Transparency

KR5. Diversity,  
non-  
discrimination  
and fairness

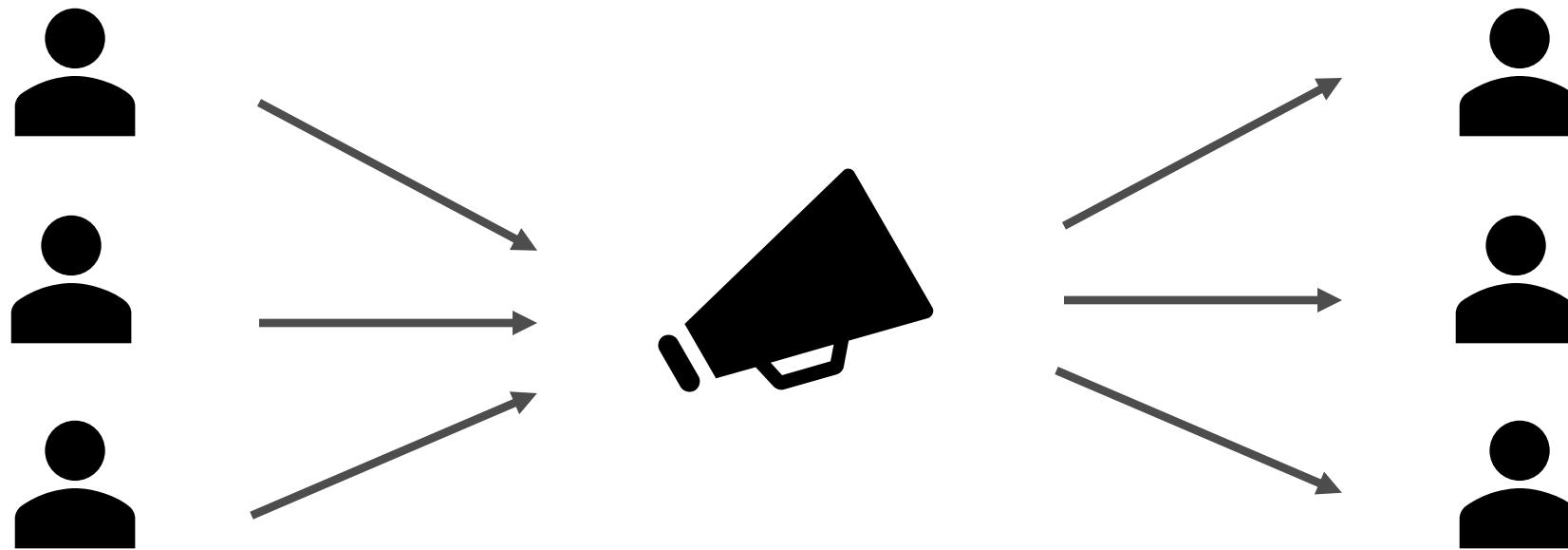
KR6. Societal and  
environmental  
well-being

KR7.  
Accountability



*“A new Digital Services Act will upgrade our liability and safety rules for digital platforms, services and products, and complete our Digital Single Market”.*

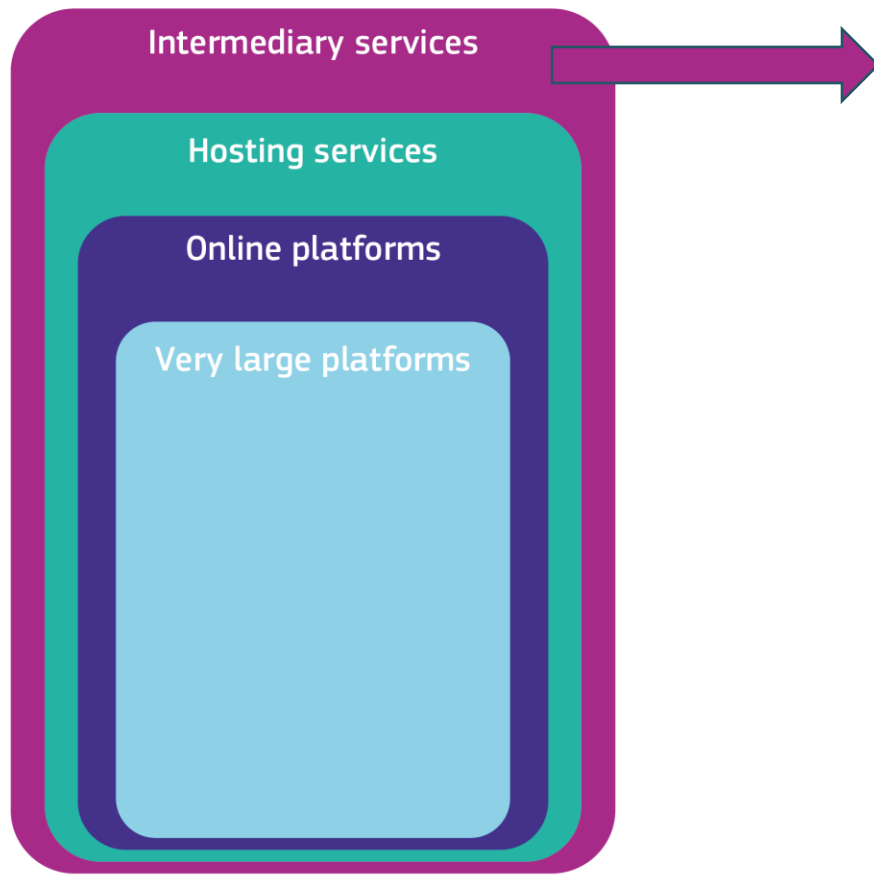
# What does the DSA regulate exactly?



**Intermediary Services**



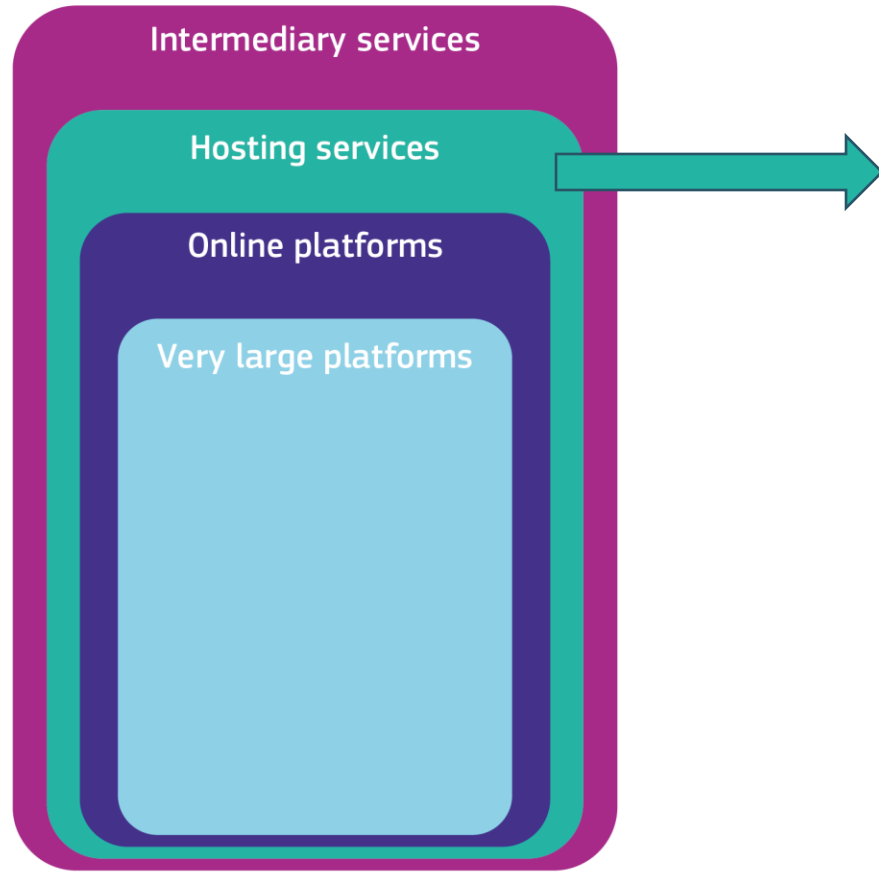
# DSA: Architecture



‘Mere conduit’: providers of communication services

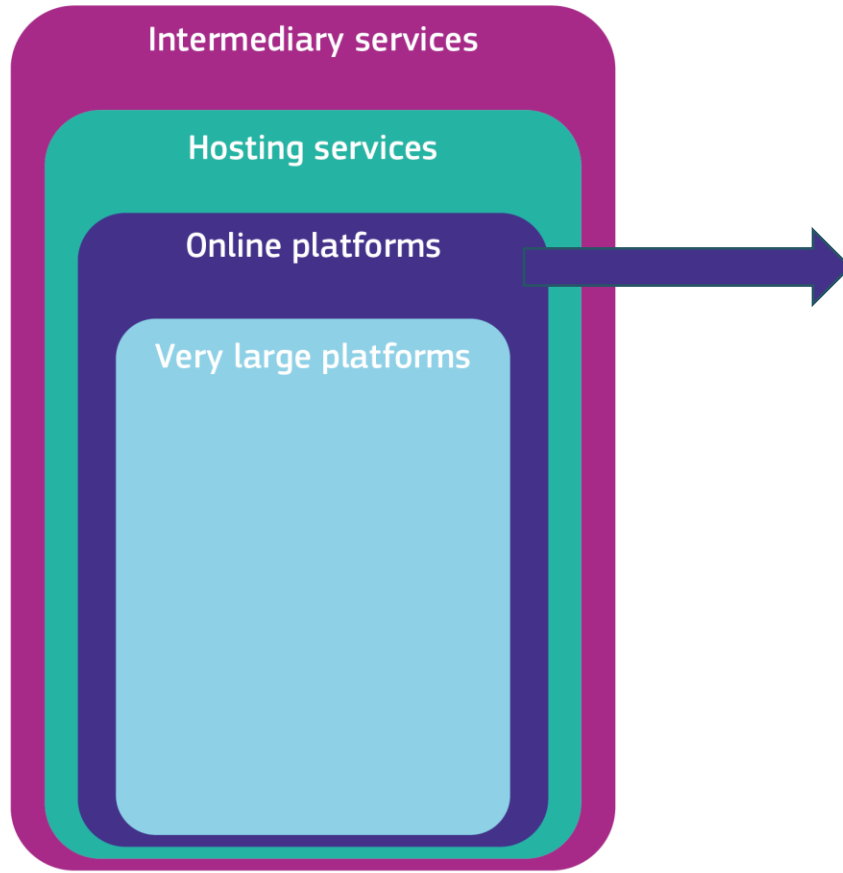
‘Caching service’: mere conduit with temporary storage for the sole purpose of making communication more efficient

# DSA: Architecture



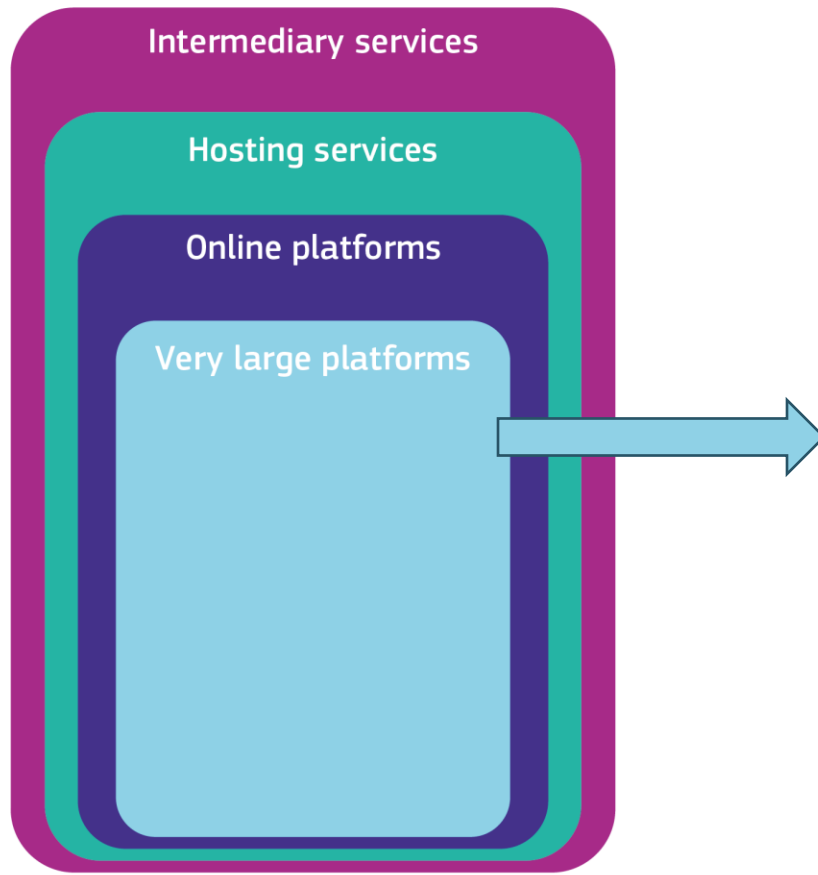
Hosting services: service for storage of information at the recipient's request

# DSA: Architecture



Online platforms: services that store and disseminate information at the recipient's request

# DSA: Architecture



Very Large Online Platforms (VLOPs) and Very Large Search Engines (VLOSEs): online platforms with more than 45 million active recipients in the EU

# DSA: Authorities



**Digital Services Coordinators (DSCs)** supervise **Intermediaries, Hosting Services and Online Platforms** in each Member State



**European Commission** supervises **VLOPs and VLOSEs** at EU level

# DSA: Designated VLOPs and VLOSEs

- AliExpress
- Amazon Store
- Apple App Store
- Booking.com
- Google Search
- Google Play
- Google Maps
- Google Shopping
- LinkedIn
- Facebook
- Instagram
- Bing
- XNXX
- Pinterest
- Pornhub
- Snapchat
- Shein
- Stripchat
- TikTok
- X
- Temu
- XVideos
- Wikipedia
- Youtube
- Zalando

# The DSA and recommender systems

**‘recommender system’ means a fully or partially automated system used by an online platform to suggest in its online interface specific information to recipients of the service or prioritise that information, including as a result of a search initiated by the recipient of the service or otherwise determining the relative order or prominence of information displayed.**

# What should recsys providers worry about?

Providers of online platforms that use recommender systems shall set out in their terms and conditions, in plain and intelligible language, **the main parameters used in their recommender systems**, as well as **any options for the recipients of the service to modify or influence** those main parameters.

- Most significant criteria used
- Reasons for the relative importance of parameters
- User customisable parameters
- + Non-profiling option on VLOPs/VLOSEs (Art 38)



# DSA Data Access (Art. 40)

For compliance checking purposes:

- Data access for DSCs and Commission - Art 40(1-3)

For the detection and understanding of systemic risks

- Data access for **vettred researchers** (private data) – Art 40(4-11)
  - Request is made to the DSC of the platform or the researcher
- Access to **publicly available data** – Art 40(12)
  - Request is made directly to platform

# DSA Data Access – conditions for vetted researcher access

- Affiliation to research organisation
- Independence from commercial interests
- Funding disclosure
- Capability of fulfilling data security and confidentiality requirements
- Necessity and proportionality
- Solely for the detection and understanding of **systemic risks** and assessment mitigation measures
- Publication of results free of charge

# DSA: Systemic risks

1. Dissemination of illegal content
2. Negative effects in the exercise of fundamental rights\*
  - Human dignity
  - Respect for private and family life
  - Protection of personal data
  - Freedom of expression and information
  - Non-discrimination
  - Respect for the rights of the child
3. Negative effects on civic discourse and electoral processes, and public security
4. Negative effects on:
  - Gender-based violence
  - Protection of public health and minors
  - Serious negative consequences to personal physical and mental well-being.

\*Charter of Fundamental Rights of the European Union

# DSA: Data access for vetted researchers



# DSA Data Access – conditions for access to publicly available data – Art. 40(12)

- ~~Affiliation to research organisation~~
- Independence from commercial interests
- Funding disclosure
- Capability of fulfilling data security and confidentiality requirements
- Necessity and proportionality
- Solely for the detection and understanding of systemic risks ~~and assessment mitigation measures~~
- ~~Publication of results free of charge~~

# DSA: other research opportunities



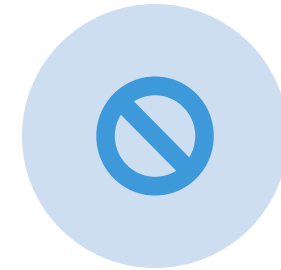
ALGORITHMIC AUDITING



EVALUATION, METRICS  
AND BENCHMARKING (E.G.  
AGAINST SYSTEMIC RISKS)



USER CUSTOMISATION  
(ART 27)



CONTENT MODERATION  
(STATEMENTS OF  
REASONS)

# Thank you

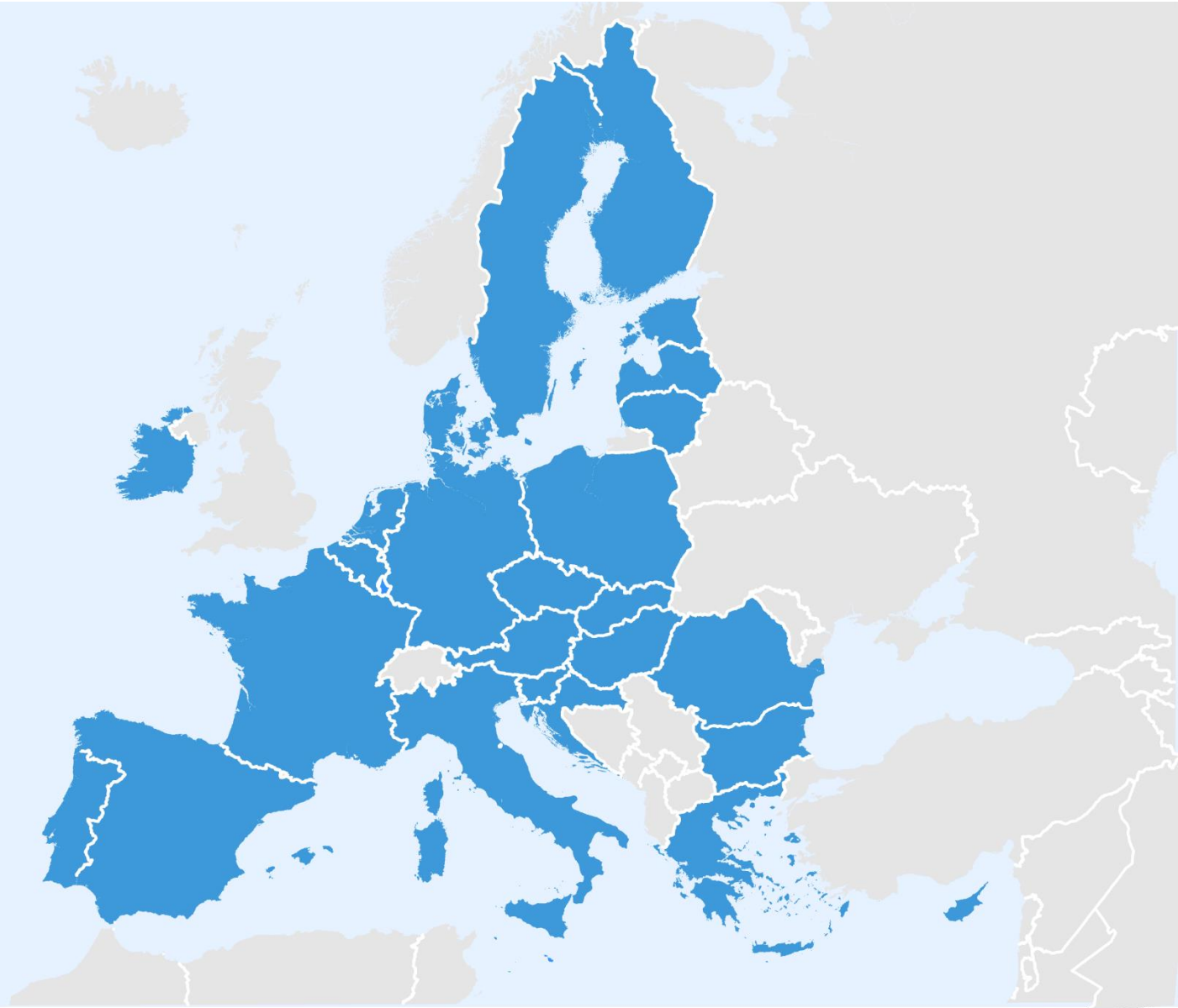
This presentation has been prepared for internal purposes. The information and views expressed in it do not necessarily reflect an official position of the European Commission or of the European Union.

Except otherwise noted, © European Union (2024). All Rights Reserved



**EU Science Hub**  
[joint-research-centre.ec.europa.eu](https://joint-research-centre.ec.europa.eu)

# EU countries



0 250 500 1,000 Km

© European Union, 2024. Map produced by EC-JRC. The boundaries and the names shown on this map do not imply official endorsement or acceptance by the European Union.