La ética de la IA en los proyectos Horizonte y las implicaciones para la investigación de la aprobación de la Ley de IA de la UE.

Lissette Lemus
lissette@iiia.csic.es
Sara Degli Esposti
sara.degli.esposti@csic.es
Pompeu Casanovas
pompeu.casanovas@iiia.csic.es

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

CSIC HUB INTELIGENCIA ARTIFICIAL

Date: 28/05/2024
12:00 - 1:30 pm

Summary

1. Objective (to furnish a description of Artificial Intelligence Act (AIA) for research purposes)
2. Vocabulary, objectives, bullet points
3. Legal Schema
4. Artificial Intelligence Act
    - Deployment, structure, content
    - Doctrine
    - Main legal principles
    - Main concepts and procedures
    - Procedures (guidelines for research)
5. Ethical Appraisal Procedure

Vocabulary

1. Privacy
2. Deployer
3. Placing on the market
4. Making available on the market
5. Reasonably foreseeable misuse
6. Notifying authority
7. Conformity assessment
8. Recall of an AI system
9. *Conformité Européene* (CE marking)
10. Post-market monitoring system
11. Market surveillance authority
12. Harmonised standard
13. Sensitive operational data
14. Post remote biometric identification system
15. Law enforcement authority (LEA)
16. Law enforcement
17. Plan-world testing plan
18. Sandbox plan
19. Regulatory sandbox
20. Profiling
21. Deep-fake
22. Widespread infringement
23. General-purpose AI model
24. High-impact capabilities
25. Systemic risk
26. General-purpose AI system
27. Floating-point operation
28. Downstream operator

Equivalencias en lenguaje jurídico

1. Intimidad
2. Responsable del despliegue
3. Introducción en el mercado
4. Comercialización
5. Uso indebido razonablemente previsible
6. Autoridad notificante
7. Evaluación de la conformidad
8. Recuperación de un sistema de IA
9. Marcado CE
10. Sistema de vigilancia poscomercialización
11. Autoridad de vigilancia del mercado
12. Norma armonizada
13. Datos operativos sensibles
14. Sistema de identificación biométrica remota en diferido
15. Autoridad garante del cumplimiento del Derecho
16. Garantía del cumplimiento del Derecho
17. Plan de la prueba en condiciones reales
18. Plan del espacio controlado de pruebas
19. Espacio controlado de pruebas para la IA
20. Elaboración de perfiles
21. Ultrasuplantación
22. Infracción generalizada
23. Modelo de IA de uso general
24. Capacidades de gran impacto
25. Riesgo sistémico
26. Sistema de IA de uso general
27. Operación de coma flotante
28. Operador posterior

Bullet points to remember (AI Act)

- AI value chain of the AI system (providers, deployers, importers, distributors and product manufacturers) (*cadena de valor de la IA*)
- AI Categories (General-purpose AI model; High-impact capabilities; Systemic risk; General-purpose AI system)
- Evaluation and management of risk levels (*niveles de riesgo*)
- Conformity assessment (*evaluación de conformidad*)
- CE Marking (*Marcado CE*)
- Codes of practice (*Códigos de buenas prácticas*)
- Regulatory sandboxes (*espacios controlados de prueba*)
- AI Office, Committee, Forum (*Oficina de AI, Comité y Foro*)

Recital 101: AI value chain

“Providers of general-purpose AI models have a particular role and responsibility along the AI value chain, as the models they provide may form the basis for a range of downstream systems, often provided by downstream providers that necessitate a good understanding of the models and their capabilities, both to enable the integration of such models into their products, and to fulfil their obligations under this or other regulations. Therefore, proportionate transparency measures should be laid down, including the drawing up and keeping up to date of documentation, and the provision of information on the general-purpose AI model for its usage by the downstream providers. Technical documentation should be prepared and kept up to date by the general-purpose AI model provider for the purpose of making it available, upon request, to the AI Office and the national competent authorities. The minimal set of elements to be included in such documentation should be set out in annexes to this Regulation. The Commission should be empowered to amend those annexes by means of delegated acts in the light of evolving technological developments.

Considerando 101

> Los proveedores de modelos de IA de uso general tienen una función y una responsabilidad particulares a lo largo de la cadena de valor de la IA, ya que los modelos que suministran pueden constituir la base de diversos sistemas de etapas posteriores, que a menudo son suministrados por proveedores posteriores que necesitan entender bien los modelos y sus capacidades, tanto para permitir la integración de dichos modelos en sus productos como para cumplir sus obligaciones en virtud del presente Reglamento o de otros reglamentos. Por consiguiente, deben establecerse medidas de transparencia proporcionadas, lo que incluye elaborar documentación y mantenerla actualizada y facilitar información sobre el modelo de IA de uso general para su uso por parte de los proveedores posteriores. El proveedor del modelo de IA de uso general debe elaborar y mantener actualizada la documentación técnica con el fin de ponerla a disposición, previa solicitud, de la Oficina de IA y de las autoridades nacionales competentes

Recital 1

Too many objectives!

> The purpose of this Regulation is to improve the functioning of the internal market by laying down a uniform legal framework in particular for the development, the *placing on the market, the putting into service and the* use of artificial intelligence *systems (*AI systems*) in the Union,* in accordance with Union values*, to promote the uptake of human centric and trustworthy artificial intelligence (AI) while ensuring* a high level of protection of health, safety, fundamental rights as *enshrined in the Charter of fundamental rights of the European Union (the 'Charter'), including democracy, the rule of law and environmental protection, against the harmful effects of AI systems in the Union, and to support innovation. This Regulation* ensures the free movement, crossborder, of AI-based goods and services, thus preventing Member States from imposing restrictions on the development, marketing and use of *AI systems*, unless explicitly authorised by this Regulation.

Article 5.1

Counterfactuals!

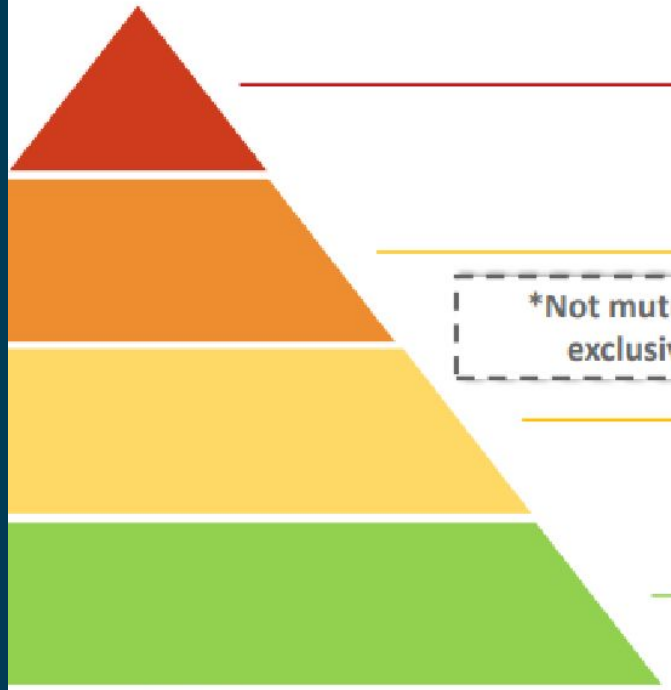
“ The following AI practices shall be prohibited:

(a) the placing on the market, the putting into service or the use of an AI system that deploys subliminal techniques beyond a person’s consciousness *or purposefully manipulative or deceptive techniques, with the objective, or the effect of,* materially *distorting* the behaviour of a person *or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing a person to take a decision that that person would not have otherwise taken* in a manner that causes or
is likely to cause that person, another person *or group of persons significant* harm;

Bear in mind:

COMPAS
Cambridge Analytica
Social Scoring Systems...

+ vulnerabilities; human rights; physical, psychological, collective harm, social scoring, biomètric identification and classification...



*Not mutually exclusive

Unacceptable risk
e.g. social scoring
— **Prohibited**

High risk
e.g. recruitment, medical devices
— **Permitted** subject to compliance with AI requirements and ex-ante conformity assessment

AI with specific transparency obligations
'Impersonation' (bots)
— **Permitted** but subject to information/transparency Obligations

Minimal or no risk
— **Permitted** with no restrictions

European Commission

How to proceed to be AIA compliant (research plan)

- Identify the EU Common Data Space in which the project is situated (e.g. Health, Finances..)
- Identify the position of the project (or some parts of the project) using the AI Act classification and categories (in the AI value chain)
- Set an Independent Ethical Committee
- Elaborate a detailed time-table (specifying AI modules, tasks, workflows, KPIs and possible or prospective entry into the market)
- Identify human decisions and behaviour that would require informed consent
- Identify possible biases (avoid discrimination)
- Elaborate a risk assessment, management and a redress plan
- Plan a dynamic data/FHR impact assessments
- Create an internal (ethical and legal) regulatory sandbox all along the development of the project (especially for pilots and use cases)
- Plan mid and final term ethical/legal internal audits
- Get an external Ethical Committee approval (high-risk or sensitive projects)
- Show compliance with Codes of Practices
- Do not trust ChatGPT!

- Industrial (manufacturing) data space, to support the competitiveness and performance of the EU's industry, allowing to capture the potential value of use of non-personal data in manufacturing (estimated at € 1,5 trillion by 2027).
- Green Deal data space, to use the major potential of data in support of the Green Deal priority actions on climate change, circular economy, zero-pollution, biodiversity, deforestation and compliance assurance.
- Mobility data space, to position Europe at the forefront of the development of an intelligent transport system, including connected cars as well as other modes of transport.
- Health data space, which is essential for advances in preventing, detecting and curing diseases as well as for informed decisions to improve the accessibility, and sustainability of the healthcare systems.
- Financial data space, to stimulate, through enhanced data sharing, innovation, market transparency, sustainable finance, as well as access to finance for European businesses and a more integrated market.
- Energy data space, to promote a stronger availability and cross-sector sharing of data,
- Agriculture data space, to enhance the sustainability performance and competitiveness of the agricultural sector through the processing and analysis of production and other data, at farm level.
- Data spaces for public administration, to improve transparency and accountability of public spending and spending quality, fighting corruption, both at EU and national level, and to address law enforcement needs and support the effective application of EU law and enable innovative 'gov tech', 'reg tech' and 'legal tech'
- Skills data space, to reduce the skills mismatches between the education and training system on the one hand and the labour market needs on the other.

Common Data Spaces (EU Strategy for Data, 2020)

What is a ‘Legal Schema’?

- Law: A ‘legal schema’ is a structured, organised and consistent summary or model of the legal provisions contained in a legal document (such as an Act, Regulation or Contract)
- AI & LawTechnology: Open source initiative that provides a common language for creating and managing legal documents as data

AI Act : Legal Schema

- The AI Act includes a dynamic complex set of procedures, permits and prohibitions, related to an EU platform and economy-driven network of regulations ("new legislation") which is difficult to turn into a simple "legal schema".
- There are pros and cons: (i) It enhances citizens' rights and stresses protections; (ii) while creating a complicated public space and a complex organisational and administrative framework, difficult to manage.
- The AIA is a EU *regulation, a specific legal instrument* defined in the *Treatise on the Functioning of the European Union* (2012), binding and directly applicable to all national states of the EU (EU jurisdiction).

REGULATION

“ The Treaty on the Functioning of the European Union (TFUE, art. 288)

To exercise the Union's competences, the institutions shall adopt regulations, directives, decisions, recommendations and opinions. A regulation shall have general application. *It shall be binding in its entirety and directly applicable in all Member States*. A directive shall be binding, as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods. *A decision shall be binding in its entirety*. A decision which specifies those to whom it is addressed shall be binding only on them. Recommendations and opinions shall have no binding force

COORDINATION (Principles: Subsidiarity-Proportionality)

“ Treaty on the Functioning of the European Union (2012, art. 5)

1. The Member States shall coordinate their economic policies within the Union. To this end, the Council shall adopt measures, in particular broad guidelines for these policies. Specific provisions shall apply to those Member States whose currency is the euro.

2. The Union shall take measures to ensure coordination of the employment policies of the Member States, in particular by defining guidelines for these policies.

3. The Union may take initiatives to ensure coordination of Member States' social policies.

“ The Treaty on the Functioning of The European Union (2012)

Article 289

1. The ordinary legislative procedure shall consist in the joint adoption by the European Parliament and the Council of a regulation, directive or decision on a proposal from the Commission. This procedure is defined in Article 294.

Article 290

1. A legislative act may delegate to the Commission the power to adopt non-legislative acts of general application to supplement or amend certain non-essential elements of the legislative act. The objectives, content, scope and duration of the delegation of power shall be explicitly defined in the legislative acts. The essential elements of an area shall be reserved for the legislative act and accordingly shall not be the subject of a delegation of power.

Deployment

- **Brussels, 21.4.2021. COM(2021) 206 final. 2021/0106(COD). Proposal** for a Regulation of the European Parliament and of the Council, laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative Acts.

- **Brussels, 13.3.2024**. P9_TA(2024)0138. **Artificial Intelligence Act**. European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)) (Ordinary legislative procedure: first reading),

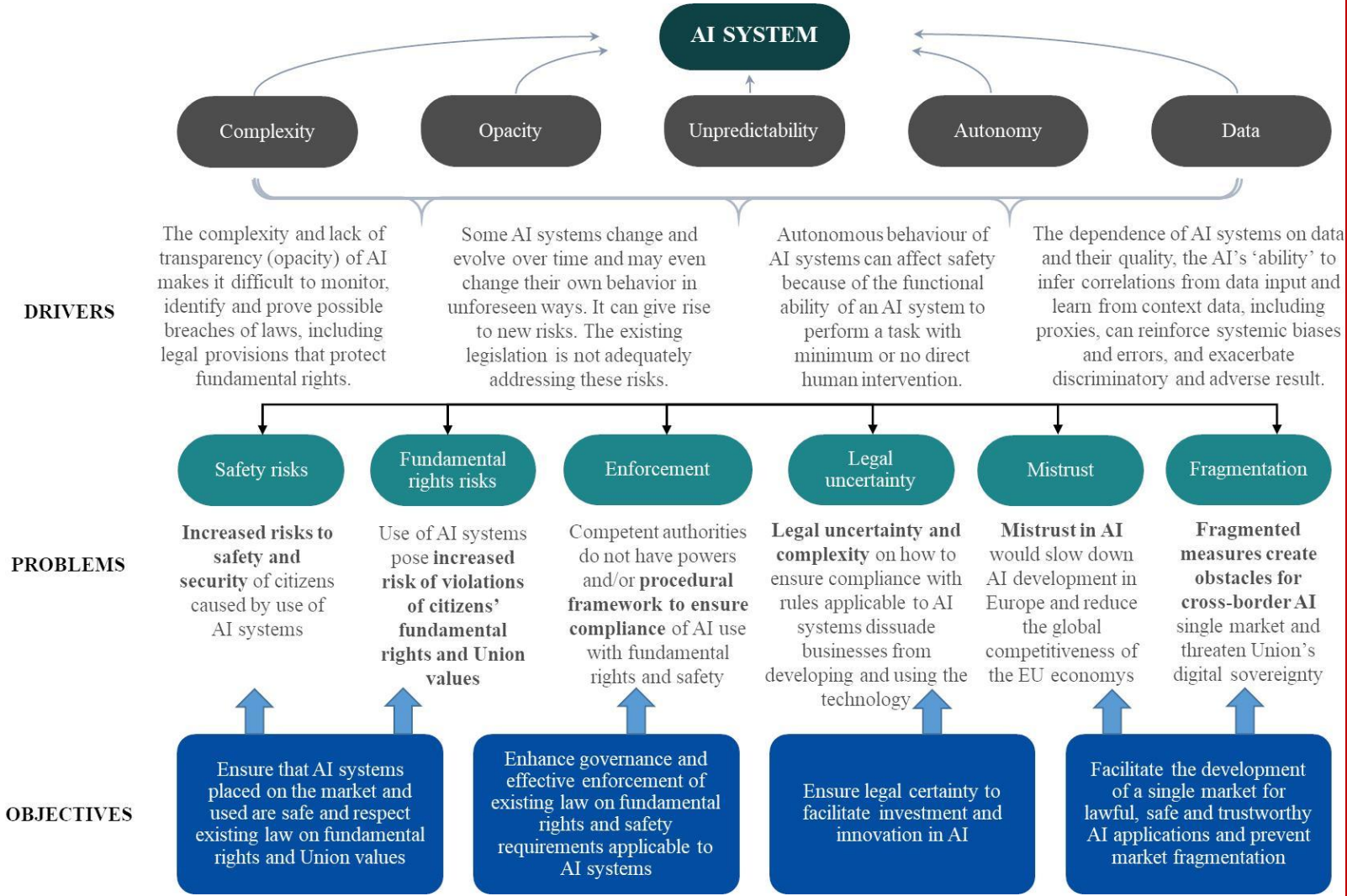
- **Bruselas, 24.5. 2024. REGLAMENTO** DEL PARLAMENTO EUROPEO Y DEL CONSEJO por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial)

Proposal (Draft 2021)

- In total, 1215 contributions were received, of which 352 were from companies or business organisations/associations, 406 from individuals (92%individuals from EU), 152 on behalf of 22 and 73 from public authorities. Civil society's voices were represented by 160 respondents (among which 9 consumers' organisations, 129 non-governmental organisations and 22 trade unions), 72 respondents contributed as 'other'. Of the 352 business and industry representatives, 222 were companies and business representatives, 41.5% of which were micro, small and medium-sized enterprises. The rest were business associations. Overall, 84% of business and industry replies came from the EU-27. Depending on the question, between 81 and 598 of the respondents used the free text option to insert comments. Over 450 position papers were submitted through the EU Survey website, either in addition to questionnaire answers (over 400) or as stand-alone contributions (over 50).

- High-Level Expert Group on AI (HLEG) which had an inclusive and broad composition of 52 well-known experts tasked to advise the Commission on the implementation of the Commission's Strategy on Artificial Intelligence

- In line with its "Better Regulation" policy, the Commission conducted an impact assessment for this proposal examined by the Commission's Regulatory Scrutiny Board



COMMISSION STAFF WORKING DOCUMENT

IMPACT ASSESSMENT Accompanying the Proposal for a Regulation of the European Parliament and of the Council

Brussels, 21.4.2021 SWD(2021) 84 final,

https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021SC0084

Final contents AIA (2024)

- Three main parts: (i) Recitals; (ii) Body (Chapters, Sections, Articles), (iii) Annexes
- More than 450 pages (Council document)
- 180 Recitals (guidelines)
- 113 Articles (content)
- 13 Annexes (amendments and procedures)

After being signed by the presidents of the European Parliament and of the Council, the legislative act will be published in the EU's Official Journal in the coming days and enter into force twenty days after this publication. The new regulation will apply two years after its entry into force, with some exceptions for specific provisions. (EU website)

The European Artificial Intelligence (AI) Act is a proposed regulation designed to ensure AI technologies in the EU are safe, transparent, and respect fundamental rights. The recitals of the Act outline its guiding principles and rationale:

1. Risk-Based Approach: AI systems are classified by risk levels—unacceptable, high, limited, and minimal. This ensures proportional regulation based on potential harm.

2. Fundamental Rights Protection: The Act emphasizes safeguarding privacy, non-discrimination, and data protection, ensuring AI systems do not infringe on fundamental rights.

3. Transparency and Accountability: AI systems must be transparent, informing users when they are interacting with AI and allowing for human oversight to mitigate risks.

4. High-Risk AI Systems: These systems are subject to strict requirements, including thorough assessments, data governance measures, and documentation to ensure safety and compliance.

5. Prohibited Practices: Certain AI applications are banned, such as social scoring by governments, exploiting vulnerabilities, and using subliminal techniques to cause harm.

6. Innovation and Competitiveness: The Act aims to foster innovation and maintain the EU's competitiveness, offering regulatory sandboxes and support for small and medium-sized enterprises (SMEs).

7. International Cooperation: It promotes global cooperation and alignment with international standards to enhance interoperability and uphold the EU's role in global AI governance.

8. Overall, the Act seeks to promote trustworthy AI development, balancing safety, fairness, and human rights with innovation and economic growth in the EU.

AIA Summary of the Recitals (ChatGPT4)

Look out! There is no legal sense in there

II AIA Body (Chapters and Sections)

Chapter I: General provisions
Chapter II: Prohibited AI practices
Chapter III: High-risk AI systems
- *Section 1: Classification of AI systems as high-risk*
- *Section 2: Requirements for high-risk AI systems*
- *Section 3: Obligations of providers and deployers of high-risk AI systems and other parties*
- *Section 4: Notifying authorities and notified bodies*
- *Section 5: Standards, conformity assessment, certificates, registration*

Chapter IV: Transparency obligations or providers and deployers of certain AI systems
Chapter V: General-purpose AI models
- *Section 1. Classification rules*
- *Section 2. Obligations for providers of general-purpose AI models*
- *Section 3. Obligations for providers of general-purpose AI models with systemic risk*

Chapter VI: Measures in support of innovation
Chapter VII: Governance.
- Section 1. Governance at Union level
- Section 2, National competent authorities

Chapter VIII: European database for high-risk AI systems
Chapter IX: Post-market Monitoring, Information Sharing, Market Surveillance
- *Section 1. Post-market monitoring*
- *Section 2. Sharing of information on serious incidents*
- *Section 3. Enforcement*
- *Section 4. Remedies*
- *Section 5. Supervision, investigation, enforcement and monitoring in respect of providers of general-purpose AI models*

Chapter X: Codes of conduct and guidelines
Chapter XI: Delegation of power and Committee procedure (art. 97)
Chapter XII: Penalties
Chapter XIII: Final Provisions:

II AIA Body (articles)

Chapter I: General provisions: subject matter (art.1) , scope-harmonisation (art.2), definitions (art. 3)
Chapter II: Prohibited AI pratices (art.5)
Chapter III: High-risk AI systems (art.6)
1. Classification (art.6)
2. Requirements and Compliance (art.8)
   Risk management system (art.9)
   Data and data governance (art.10)
   Technical documentation (art.11)
   Record-keeping (art.12)
   Transparency and provision of information to deployers (art.13)
   Human oversight (supervision) (art.14)
   Accuracy, robustness and cybersecurity (art.15)
3. Obligations of providers, deployers and other parties (art.16-27)
4. Notifying authorities and noified bodies. Notification procedure (requirements related to notified bodies) (art.28-39)
5. Standards, conformity assessment, certificates, marks, registration (art. 40-59)
Chapter IV: Transparency obligations for providers and deployers of certain AI systems (art.50)
Chapter V: General-purpose AI models. Obligations, GP-AI with systemic risk (art- 51-55), Codes of practice (art. 56)
Chapter VI: Measures in support to innovation: Regulatory sandboxes (art.57-61)
Chapter VII. Governance.
- Union level: 1.AI European AI Board , 2. Advisory Forum, 3.Scientific panel of independent experts (art. 65-69)
- National competent authorities and single point of contact (art. 70)

Chapter VIII: European database for high-risk AI systems (art. 71)
Chapter IX: Post-market Monitoring, Information Sharing, Market Surveillance , Enforcement and Control (art. 72-94)
Chapter X: Codes of conduct and guidelines (art.95-97)
Chapter XI: Delegation of power and Committee procedure (art. 97)
Chapter XII: Penalties: administrative fines on Union institutions, bodies, offices and agencies (art. 99); fines for providers of general-purpose AI models (art. 100-101)
Chapter XIII: Final Provisions: Amendments to Directives and Regulations (102-109), AI systems already placed on the market or put into service (art.110), Entry into forceand application (art.113)

II AIA Body (articles)

Chapter I: General provisions: subject matter (art.1) , scope-harmonisation (art.2), definitions (art. 3) Annex I, X **Annex II**

Chapter II: Prohibited AI pratices (art.5)

Chapter III: High-risk AI systems (art.6) Annex III

1. Classification (art.6)

2. Requirements and Compliance (art.8)
   - Risk management system (art.9)
   - Data and data governance (art.10)
   - Technical documentation (art.11) Annex IV
   - Record-keeping (art.12)
   - Transparency and provision of information to deployers (art.13)
   - Human oversight (supervision) (art.14)
   - Accuracy, robustness and cybersecurity (art.15)

3. Obligations of providers, deployers and other parties (art.16-27)

4. Notifying authorities and notified bodies. Notification procedure (requirements related to notified bodies) (art.28-39)

5. Standards, conformity assessment, certificates, marks, registration (art. 40-51) Annex V, VI, VII ,VIII, IX

Chapter IV: Transparency obligations for providers and deployers of certain AI systems (art.52)

Chapter V: General-purpose AI models. Obligations, GP-AI with systemic risk (art- 53-55), Annex XI, XII, XIII Codes of practice (art. 56)

Chapter VI: Measures in support to innovation: Regulatory sandboxes (art.57-61) Annex IX

Chapter VII. Governance.

- Union level: 1.AI European AI Board , 2. Advisory Forum, 3.Scientific panel of independent experts (art. 65-69)
- National competent authorities and single point of contact (art. 70)

Chapter VIII: European database for high-risk AI systems (art. 71)

Chapter IX: Post-market Monitoring, Information Sharing, Market Surveillance , Enforcement and Control (art. 72-94)

Chapter X: Codes of conduct and guidelines (art.95-97)

Chapter XI: Delegation of power and Committee procedure (art. 97)

Chapter XII: Penalties: administrative fines on Union institutions, bodies, offices and agencies (art. 99); fines for providers of general-purpose AI models (art. 100-101)

Chapter XIII: Final Provisions: Amendments to Directives and Regulations (102-109), AI systems already placed on the market or put into service (art.110), Entry into forcé and application (art.113)

III. AIA Annexes I-XIII

Annex I: List of Union harmonisation legislation
Annex II: List of criminal offences referred to in Article 5(1), point (e)(iii)
Annex III: High-risk AI systems referred to in Article 6(2)
Annex IV: Technical documentation referred to in Article 11(1)
Annex V: EU declaration of conformity
Annex VI: Conformity assessment procedure based on internal control
Annex VII: Conformity based on an assessment of the quality management system and an assessment of the technical documentation
Annex VIII: Information to be submitted upon the registration of high-risk AI systems in accordance with Article 49
Annex IX:
- Section A - Information to be submitted by providers of high-risk AI systems in accordance with Article 49(1)
- Section B- Information to be submitted by providers of high-risk AI systems in accordance with Article 49(2)
- Section C- Information to be submitted by deployers of high-risk AI systems in accordance with Article 49(3)
- Section D. Information to be submitted upon the registration of high-risk AI systems listed in Annex III in relation to testing in real world conditions in accordance with Article 60

Annex X: Union legislative acts on large-scale IT systems in the area of Freedom, Security and Justice
Annex XI: Technical documentation referred to in Article 53(1), point (a) - technical documentation for providers of general-purpose AI models.
- Section 1. Information to be provided by all providers of general-purpose AI models
- Section 2. Additional information to be provided by providers of general-purpose AI models with systemic risk

Annex XII: Transparency information referred to in Article 53(1), point (b) - technical documentation for providers of general-purpose AI models to downstream providers that integrate the model into their AI system
Annex XIII: Criteria for the designation of general-purpose AI models with systemic risk referred to in Article 51

- Industrial (manufacturing) data space, to support the competitiveness and performance of the EU's industry, allowing to capture the potential value of use of non-personal data in manufacturing (estimated at € 1,5 trillion by 2027).
- Green Deal data space, to use the major potential of data in support of the Green Deal priority actions on climate change, circular economy, zero-pollution, biodiversity, deforestation and compliance assurance.
- Mobility data space, to position Europe at the forefront of the development of an intelligent transport system, including connected cars as well as other modes of transport.
- Health data space, which is essential for advances in preventing, detecting and curing diseases as well as for informed decisions to improve the accessibility, and sustainability of the healthcare systems.
- Financial data space, to stimulate, through enhanced data sharing, innovation, market transparency, sustainable finance, as well as access to finance for European businesses and a more integrated market.
- Energy data space, to promote a stronger availability and cross-sector sharing of data,
- Agriculture data space, to enhance the sustainability performance and competitiveness of the agricultural sector through the processing and analysis of production and other data, at farm level.
- Data spaces for public administration, to improve transparency and accountability of public spending and spending quality, fighting corruption, both at EU and national level, and to address law enforcement needs and support the effective application of EU law and enable innovative 'gov tech', 'reg tech' and 'legal tech'
- Skills data space, to reduce the skills mismatches between the education and training system on the one hand and the labour market needs on the other.

Common Data Spaces (EU Strategy for Data, 2020)

Legal Scheme: AIA Principles (backbones)

- Subsidiarity (enhancing UE powers vs. national states)
- Proportionality (to balance EU/national positions; and different compliant interests)
- Complementarity (legal harmonisation)
- Legal security (certainty of legal provisions and rules)

Recital 176: EU proactivity

> Since the objective of this Regulation, namely to improve the functioning of the internal market and promoting the uptake of human centric and trustworthy AI, while ensuring a high level of protection of health, safety, fundamental rights enshrined in the Charter, including democracy, the rule of law and environmental protection against harmful effects of AI systems in the Union and supporting innovation, cannot be sufficiently achieved by the Member States and can rather, by reason of the scale or effects of the action, be better achieved at Union level, the Union may adopt measures in accordance with the principle of subsidiarity as set out in Article 5 TEU. In accordance with the principle of proportionality as set out in that Article, this Regulation does not go beyond what is necessary in order to achieve that objective.

REACTIONS

1. Legal scholarship: “Regulatory brutality” (de Hert)

- We can identify three basic phenomena common to all, or most, EU new technology-relevant regulatory initiatives, namely (a) “act-ification”, (b) “GDPR mimesis”, and (c) “regulatory brutality”. These phenomena divulge new-found confidence on the part of the EU technology legislator, who has by now asserted for itself the right to form policy options and create new rules in the field for all of Europe. These three phenomena serve as indicators or early signs of a new European technology law-making paradigm that by now seems ready to emerge

- ‘Regulatory brutality’ denotes the almost complete disregard by the EU law-maker of Member States’ own legal systems while regulating technology (new terms, new procedures, new principles, new state mechanisms)

Papakonstantinou, V., & De Hert, P. (2022). The Regulation of Digital Technologies in the EU:: The law-making phenomena of “act-ification”,“GDPR mimesis” and “EU law brutality”. Technology and Regulation, 2022, 48-60.

REACTIONS

2. Ethical scholarship. *Dura Lex, sed Lex* (Floridi)

- The era of self-regulation, as a strategy for dealing with the ethical challenges posed by the digital revolution, is over

- The time has come to acknowledge that, much as it might have been worth trying, self-regulation did not work. So, to use the words of the Gospel, now that the invitation has not been accepted, the alternative is "to force them [companies] to enter" (Luke 14:23).

- Self-regulation needs to be replaced by the law; the sooner, the better. *Dura lex, sed lex* digitalis is why the EU is at the forefront in the debate on digital governance

L. Floridi: Dura lex sed lex "The End of an Era: from Self-Regulation to Hard Law for the Digital Industry" *Philosophy & Technology* 34, no. 4 (2021): 619-622.

REACTIONS

3. Legal scholarship: A "patchwork effect" (Pagallo)

Scholars have widely discussed whether the Artificial Intelligence (AI) Act of EU law will trigger a new Brussels effect, namely, the power that EU law exerts beyond its own boundaries and jurisdiction in such fields of regulation as data protection, environmental law, or antitrust. The paper argues that both exogenous and endogenous reasons suggest that this will not be the case with the AI Act. In addition to competition among legal systems, e.g., U.S. law, several crucial limits in the normative design of the AI Act support the claim. The result will rather be a patchwork effect. On the one hand, some pieces of the EU legislation on bans of technology and high-risk uses of AI can affect other jurisdictions and the private sector; yet, on the other hand, the troubles with the overall architecture of the regulation will make it unexportable with all its dichotomies and corresponding drawbacks.

Ugo Pagallo, "Why the AI Act Won't Trigger a Brussels Effect", AICOL2023, LNAI , Springer (in press)

AIA Analysis

Art. 1. Subject matter

“ 1. The purpose of this Regulation is to improve the functioning of the internal market and promote the uptake of human-centric and trustworthy artificial intelligence (AI), while ensuring a high level of protection of health, safety, fundamental rights enshrined in the Charter of Fundamental Rights, including democracy, the rule of law and environmental protection, against the harmful effects of artificial intelligence systems (AI systems) in the Union, and to support innovation.

Art. 1. Subject matter

“ 1.The purpose of this Regulation is to improve the functioning of the internal market and promote the uptake of human-centric and trustworthy artificial intelligence (AI), while ensuring a high level of protection of health, safety, fundamental rights enshrined in the Charter of Fundamental Rights, including democracy, the rule of law and environmental protection, against the harmful effects of artificial intelligence systems (AI systems) in the Union, and to support innovation.

Framework!

Tension!

Art. 1. Subject matter

“

2. This Regulation lays down:

(a) harmonised rules for the placing on the market, the putting into service, and the use of AI systems in the Union;
(b) prohibitions of certain AI practices;
(c) specific requirements for high-risk AI systems and obligations for operators of such systems;
(d) harmonised transparency rules for *certain* AI systems*;*
(e) harmonised rules for the placing on the market of general-purpose AI models;
(f) rules on market monitoring, *market surveillance governance and enforcement;*
(g) measures to support innovation, with a particular focus on SMEs, including startups.

LEGAL SCHEMA

Reactive
Market-driven
Systemic
Structural
General
Procedural
Hetero-regulatory
Mandatory
(top-down)
“Ethical”
Controversial

Art. 3. Definitions

“ (1) ‘AI system’ means 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 the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments;

(2) ‘risk’ means the combination of the probability of an occurrence of harm and the severity of that harm;

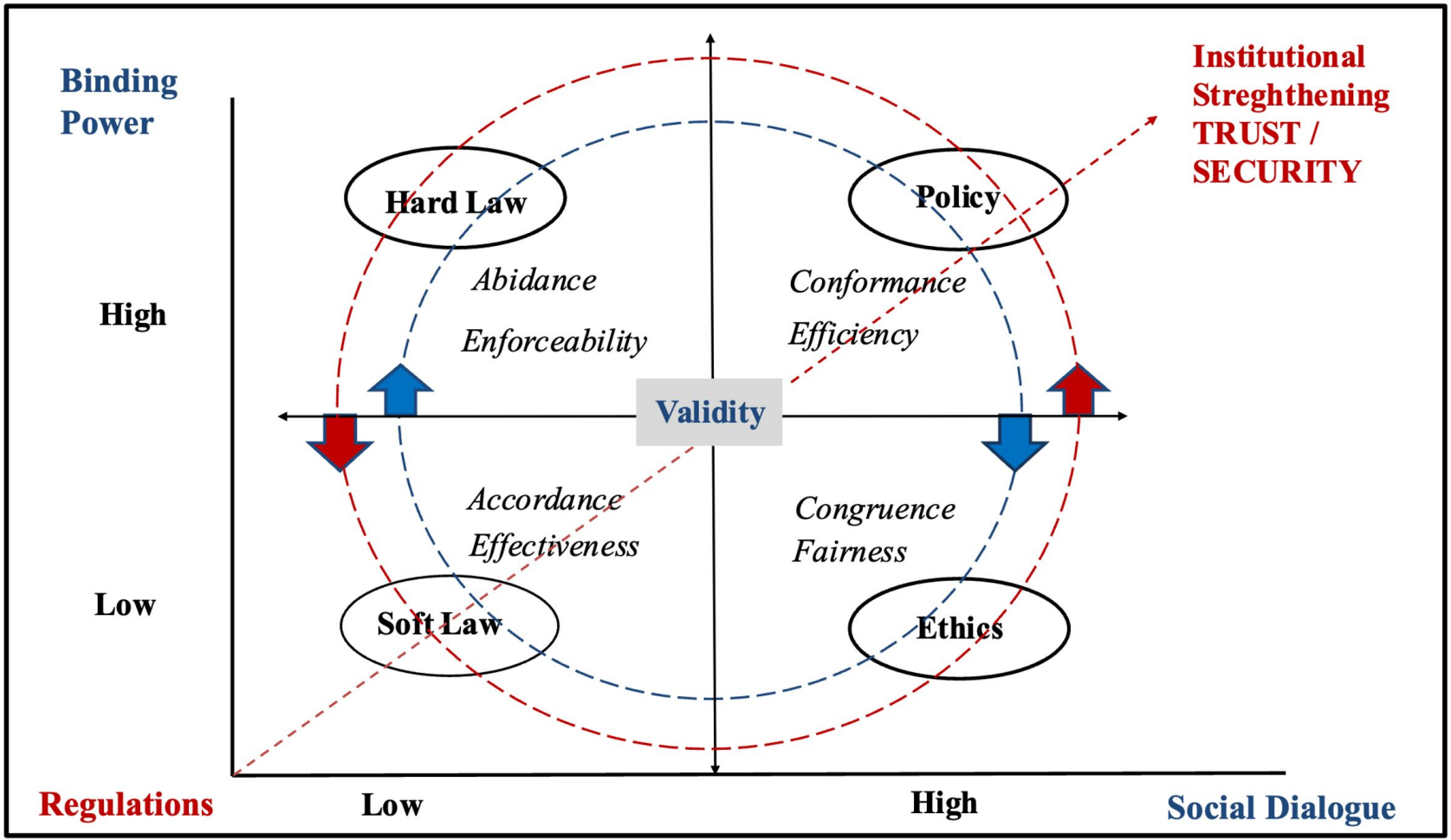
An AI system is a machine-based system that, for explicit or explicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. ***Different AI systems*** **vary** ***in their levels of autonomy and adaptiveness after deployment.***

COMPARISON to OECD ongoing definition (2023)

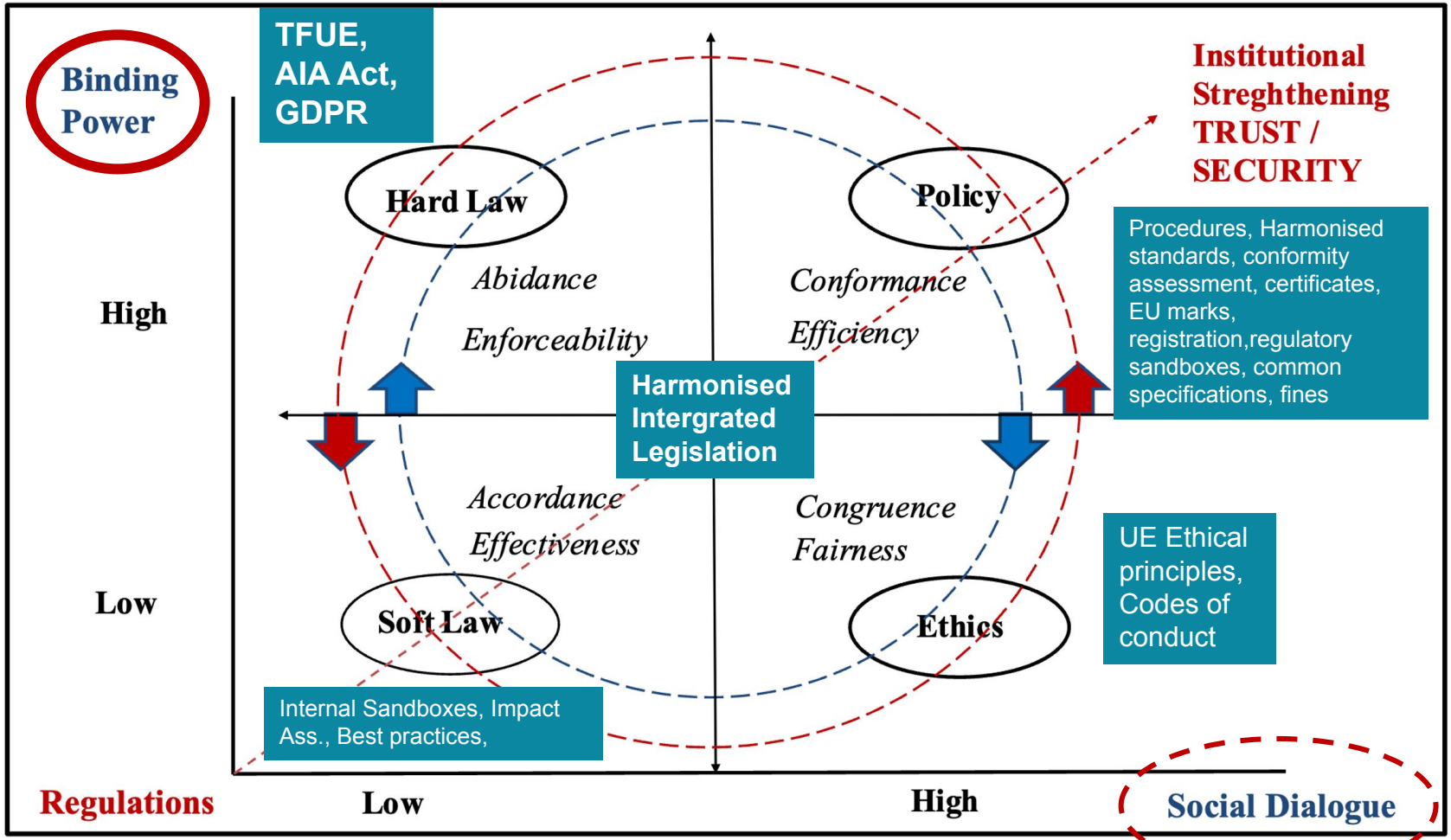
https://oecd.ai/en/wonk/definition

This tension emerges in the layout of the legal schema

- No specific risk scenarios nor metrics are provided to allocate, test and evaluate risks along the AI supply chain
- This is unclear how to separate *exceptions* from *exclusions*, even if some agents are deemed to be out of the scope of the Regulation (this is particularly important for calibrating the legal liability of researchers, who could be fined or sanctioned under some special circumstances)
- Thus, some rules contain "contrary-to-duty" obligations (deòntic ambiguity) and open-texture expressions. This indicates a propensity for stark compliance and extensive interpretation.
- Some rules show the aim to obtain general compliance (rather than conformance) with the obligations set by the AIA.
- Thus, AIA instruments (national and Eu acts, policies,ethics and soft law) are "coordinated" under the harmonised and integrated EU binding power (which the AIA delegates to the Commission and EU agents)
- AI "operators" are bound (although AIA provisions say "encouraged") to cooperate ("willingly"), according to the risk scale set by the AIA.
- Independent experts have the obligation to inform as soon as they know about a possible AIA model that does not comply with the AIA provisions



Legal Compass: Casanovas, P.; Hashmi, M., de Koker, L. (2022, 2023)



Casanovas, P.; Hashmi, M., de Koker, L. (2022, 2023)

The lawful 'AI supply chain'

Subjects:
- Providers
- Deployers
- Authorised representativ.
- Importers
- Distributors
- Downstream providers

Operators
- AI office
- Not. auth.
- National autor.,
- LEAs

Prohibitions:
- Deception, Manipulation
- Phys. Psych. Harm
- Biometrics

Criminal Offences:
- Terrorism,
- Trafficking in humans,
- Children exploitation narcotics
- Weapons
- Murder
- Human organs
- Nuclear traficking
- Kidnapping
- Hostage-taking

Classification:
- General-purpose AI models (GPAIM)
- GPAIM with systemic risks
- GPAIM with high impact capabilities
- GP AI Systems

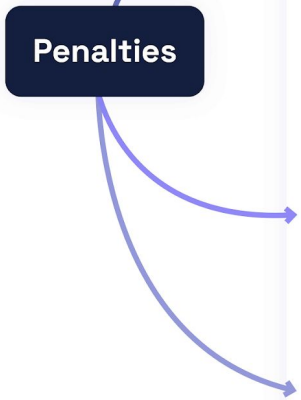
High-risk Systems::
- Biometrics
- Critical infrastructures
- Educational and vocational training
- Employment, workers management and access to self-employment
- Access to essential private services and public services and benefits
- Law enforcement
- Migration, asylum and border control management,
- Administration of justice and democratic processes

Instruments
- Conformity assessments
- CE marking
- Post-market monitoring system
- Harmonised standards
- HFR impact assessments
- Common specifications
- AI regulatory sandboxes

Procedures (Annexes):
- Technical documentation referred to in Article 11(1)
- Conformity assessment and conformity ass. based on internal control
- EU Declaration of conformity
- Ass. of the quality manag. system
- Registration of H-R systems (providers, deployers…operators) 49 (1-4)
- Sandboxes
- Testing in real world conditions (60)
- General-purpose AI models (53)
- AI models with systemic risk
- Downstream providers that integrate the model into their AI system

Exclusions (or exceptions):
Personal use
Models and systems for the sole purpose of SR&D
Or prior to their being placed on the market or put into service,
Or for military, defence or national security purposes
Or systems released under free and open source licences, unless they are placed on the market or put into service as high-risk AI systems

Penalties (Commission, Nat. surveillance market auth, Data Pr. Off. and to all operators)
For public entities
For operators (warnings and fines)



General fines for operators of AI systems	**Up to 35 000 000 EUR** or, if the offender is a company, up to 7 % of its total worldwide annual turnover for the preceding financial year, whichever is higher	Non-compliance with the prohibition of the artificial intelligence practices under Article 5
	Up to 15 000 000 EUR or, if the offender is a company, up to 3% of its total worldwide annual turnover for the preceding financial year, whichever is higher	Non-compliance with the following: • obligations of providers pursuant to Article 16 • obligations of authorised representatives pursuant to Article 25 • obligations of importers pursuant to Article 26 • obligations of distributors pursuant to Article 27 • obligations of deployers pursuant to Article 29, paragraphs 1 to 6a • requirements and obligations of notified bodies pursuant to Article 33, 34(1), 34(3), 34(4), 34a • transparency obligations for providers and users pursuant to Article 52
	Up to 7 500 000 EUR or, if the offender is a company, up to 1 % of its total worldwide annual turnover for the preceding financial year, whichever is higher	The supply of incorrect, incomplete or misleading information to notified bodies and national competent authorities in reply to a request
General fines for operators of AI systems	**Up to 15 000 000 EUR** or, if the offender is a company, up to 3% of its total worldwide annual turnover for the preceding financial year, whichever is higher	In one of the following: • Infringement of the GPAI-relevant provisions • Failure to comply with a request for document or information pursuant to Article 68i or supply of incorrect, incomplete or misleading information • Failure to comply with a measure requested under Article 68k • Failure to make available to the Commission access to the general purpose AI model or general purpose AI model with systemic risk with a view to conduct an evaluation pursuant to Article 68j
For Union institutions, agencies and bodies	**Up to EUR 1 500 000**	Non-compliance with the prohibition of the artificial intelligence practices under Article 5
	Up to EUR 750 000	Non-compliance of the AI system with any requirements or obligations

Art. 99 Penalties

“ *Non-compliance with the prohibition of the AI practices referred to in Article 5* shall be subject to administrative fines of up to *35 000 000* EUR or, if the offender is *an undertaking*, up to *7* % of its total worldwide annual turnover for the preceding financial year, whichever is higher.

4.Non-compliance of *an* AI system with any *of the following provisions related to operators or notified bodies*, other than those laid down in Articles 5, shall be subject to administrative fines of up to *15 000 000* EUR or, if the offender is an undertaking, up to *3 %* of its total worldwide annual turnover for the preceding financial year, whichever is higher. (,,,)

The supply of incorrect, incomplete or misleading information to notified bodies or national competent authorities in reply to a request shall be subject to administrative fines of up to 7 500 000 EUR or, if the offender is an undertaking, up to 1 % of its total worldwide annual turnover for the preceding financial year, whichever is higher.

Art. 3. Definitions (Agents), *within the 'AI supply chain'*

(3) 'provider' means a natural or legal person, public authority, agency or other body that develops an AI system or *a general-purpose AI model or* that has an AI system *or* a general-purpose AI model developed and *places it on the market* or puts the AI system into service under its own name or trademark, whether for payment or free of charge
(4) 'deployer' means a natural or legal person (…) using an AI system (…) except where the AI system is used in the course of a personal non-professional activity;
(5) 'authorised representative' means a natural or legal person (…) who has received and accepted a written mandate from a provider of an *AI system* or a *general-purpose AI model* to (…) carry out the obligations and procedures established by this Regulation;
(6) 'importer' means a natural or legal person located or established in the Union that places on the market an AI system (…)
(7) 'distributor' means a natural or legal person in the supply chain (…) the importer, that makes an AI system available on the Union market

Recital 178: *Providers of high-risk AI systems are encouraged to start to comply, on a voluntary basis, with the relevant obligations of this Regulation already during the transitional period*.

All of them (3-7) are (8) 'operators' with a set of different tasks and obligations, BUT they can be equally deemed legally liable under the EU AIA, as laid down by art. 21.5:

Art. 25.1 Artificial IA

“ Responsibilities along the AI value chain

1. Any distributor, importer, deployer or other third-party shall be considered to be a provider of a high-risk AI system for the purposes of this Regulation and shall be subject to the obligations of the provider under Article 16, in any of the following circumstances:

(a) they put their name or trademark on a high-risk AI system already placed on the market or put into service, without prejudice to contractual arrangements
stipulating that the obligations therein are allocated otherwise;

(b) they make a substantial modification to a high-risk AI system that has already been placed on the market or has already been put into service in such a way that it remains a high-risk AI system pursuant to Article 6;

(c) they modify the intended purpose of an AI system, including a general-purpose AI system, which has not been classified as high-risk and has already been placed on the market or put into service in such a way that the AI system concerned becomes a high-risk AI system in accordance with Article 6.

Recital 109 (deontic ambiguity, open texture)

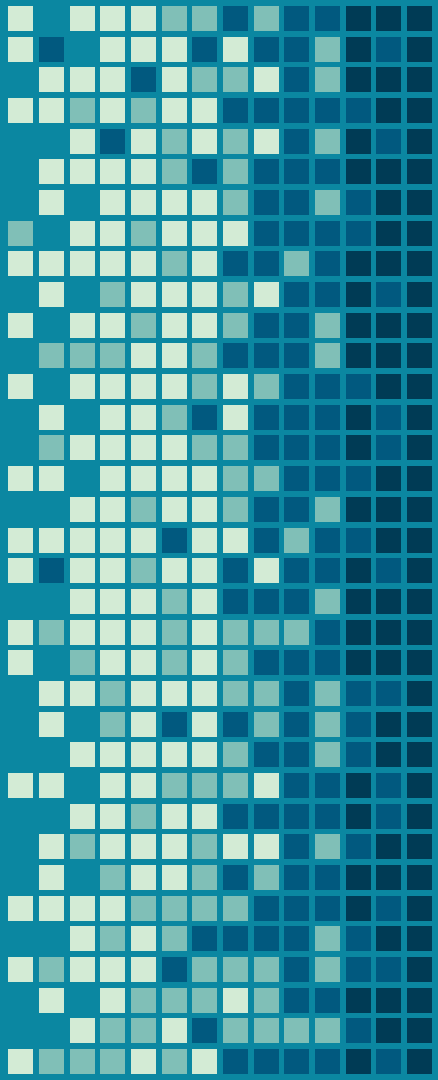
“Compliance with the obligations applicable to the providers of general-purpose AI models should be commensurate and proportionate to the type of model provider, excluding the need for compliance for persons who develop or use models for non professional or scientific research purposes, who should nevertheless be encouraged to voluntarily comply with these requirements. Without prejudice to Union copyright law, compliance with these obligations should take due account of the size of the provider and allow simplified ways of compliance for SMEs, including start-ups, that should not represent an excessive cost and not discourage the use of such models. In the case of a modification or fine-tuning of a model, the obligations for providers should be limited to that modification or fine-tuning, for example by complementing the already existing technical documentation with information on the modifications, including new training data sources, as a means to comply with the value chain obligations provided in this Regulation.

Recital 163 (suspicion, alerts, investigatory power)

With a view to complement the governance systems for general-purpose AI models, the scientific panel should support the monitoring activities of the AI Office and may, in certain cases, provide qualified alerts to the AI Office which trigger follow-ups such as investigations. This should be the case where the scientific panel has reason to suspect that a general-purpose AI model poses a concrete and identifiable risk at Union level. Furthermore, this should be the case where the scientific panel has reason to suspect that a general-purpose AI model meets the criteria that would lead to a classification as general-purpose AI model with systemic risk. To equip the scientific panel with the information necessary for the performance of those tasks, there should be a mechanism whereby the scientific panel can request the Commission to require documentation or information from a provider.

Recital 174 (extension to all AI systems)

> *Finally, by … [four years from the entry into force of this Regulation] and every three years thereafter, the Commission should evaluate the impact and effectiveness of voluntary codes of conduct to foster the application of the requirements provided for high-risk AI systems in the case of AI systems other than high-risk AI systems and possibly other additional requirements for such AI systems*



Article 59, 2 (criminal liability in regulatory sandboxes)

> *For the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including safeguarding against and preventing prevention threats to public security, under the control and responsibility of law enforcement authorities, the processing of personal data in AI regulatory sandboxes shall be based on a specific or Union or national law and subject to the same cumulative conditions as referred to in paragraph 1.*

1.c. there are effective monitoring mechanisms to identify if any high risks to the *rights and freedoms* of the data subjects, *as referred to in Article 35 of Regulation (EU) 2016/679 and in Article 39 of Regulation (EU) 2018/1725,* may arise during the sandbox experimentation, as well as response mechanisms to promptly mitigate those risks and, where necessary, stop the processing

How to proceed to be AIA compliant (research plan)

- Identify the EU Common Data Space in which the project is situated (e.g. Health, Finances..)
- Identify the position of the project (or some parts of the project) using AIA classification and categories
- Set an Independent Ethical Committee
- Elaborate a detailed time-table (specifying AI modules, tasks, workflows, and possible commodification or entry into the market)
- Identify human decisions and behaviour that would require informed consent
- Identify possible biases (avoid discrimination)
- Elaborate a risk assessment and a redress plan
- Plan a dynamic data/FHR impact assessment
- Create an internal (ethical and legal) regulatory sandbox all along the development of the project (especially for pilots and use cases)
- Plan mid and final term ethical/legal internal audits
- Get an external Ethical Committee approval (high-risk or sensitive projects)
- Do not trust ChatGPT!

AI Ethics in Horizon projects and the implications for research of the approval of the EU AI Act
May 28 2024, 12:00 - 13:30, online.

AI in Horizon Europe Ethics Screening

Sara degli Esposti – sara.degli.esposti@csic.es

Until 2027

Horizon Europe

- Horizon Europe is the EU's key funding programme for research and innovation with a budget of €95.5 billion.

Horizon Europe

- Horizon Europe is the EU's key funding programme for research and innovation with a budget of €95.5 billion.

Horizon Europe

- Horizon Europe is the EU's key funding programme for research and innovation with a budget of €95.5 billion.

Cluster 1	Health	€8.246 billion (including €1.35 billion from NGEU)
Cluster 2	Culture, Creativity & Inclusive Societies	€2.280 billion
Cluster 3	Civil Security for Society	€1.596 billion
Cluster 4	Digital, Industry & Space	€15.349 billion (including €1.35 billion from NGEU)
Cluster 5	Climate, Energy & Mobility	€15.123 billion (including €1.35 billion from NGEU)
Cluster 6	Food, Bioeconomy, Natural Resources, Agriculture & Environment	€8.952 billion
	JRC (non-nuclear direct actions)	€1.970 billion

Horizon Europe

- Horizon Europe is the EU's key funding programme for research and innovation with a budget of €95.5 billion.

Pillar III

INNOVATIVE EUROPE:

stimulating **market-creating breakthroughs** and **ecosystems** conducive to innovation

European Innovation Council

Support to innovations with breakthrough and market creating potential

European innovation ecosystems

Connecting with regional and national innovation actors

European Institute of Innovation and Technology (EIT)

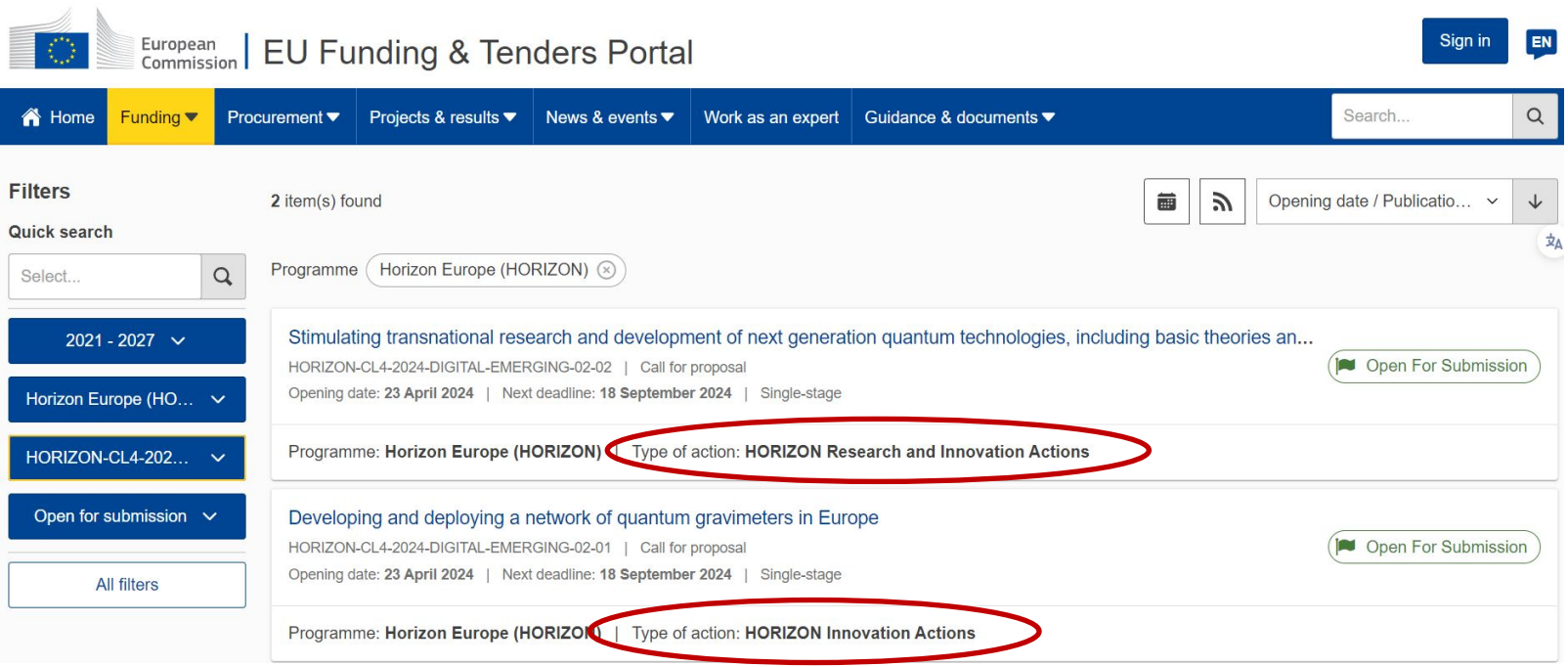
Bringing key actors (research, education and business) together around a common goal for nurturing innovation

The budget: **€10.6 billion**, incl. up to **€527 million** for ecosystems (including NGEU – Recovery Fund parts dedicated to EIC).

circa **€3 billion**

EU Funding & Tenders Portal

- Partner search
- Online Manual
  - Search funding opportunities — Find a call
- Horizon Europe (HORIZON) / Calls for proposals
- ERC Ethics guidance
- Ethics in Horizon Europe (new webpage coming soon)


European Commission | EU Funding & Tenders Portal
Sign in
EN
Home
Funding
Procurement
Projects & results
News & events
Work as an expert
Guidance & documents
Search...
Filters
Quick search
Select...
2021 - 2027
Horizon Europe (HO...
HORIZON-CL4-202...
Open for submission
All filters
2 item(s) found
Opening date / Publicatio...
Programme
Horizon Europe (HORIZON)
Stimulating transnational research and development of next generation quantum technologies, including basic theories an...
HORIZON-CL4-2024-DIGITAL-EMERGING-02-02 | Call for proposal
Opening date: 23 April 2024 | Next deadline: 18 September 2024 | Single-stage
Open For Submission
Programme: Horizon Europe (HORIZON) | Type of action: HORIZON Research and Innovation Actions
Developing and deploying a network of quantum gravimeters in Europe
HORIZON-CL4-2024-DIGITAL-EMERGING-02-01 | Call for proposal
Opening date: 23 April 2024 | Next deadline: 18 September 2024 | Single-stage
Open For Submission
Programme: Horizon Europe (HORIZON) | Type of action: HORIZON Innovation Actions

Application forms / PART A

Table of contents

4 - Other questions

Ethics Issues Table

1. Human embryonic stem cells and human embryos			Page
Does this activity involve human embryonic stem cells (hESCs)?	○ Yes	◉ No	
Does this activity involve the use of human embryos?	○ Yes	◉ No	
2. Humans			**Page**
Does this activity involve human participants?	○ Yes	◉ No	
Does this activity involve interventions (physical also including imaging technology, behavioural treatments, tracking and tracing etc.) on the study participants?	○ Yes	◉ No	
Does this activity involve conducting a clinical study as defined by the Clinical Trial Regulation (EU 536/2014) (using pharmaceuticals, biologicals, radiopharmaceuticals, or advanced therapy medicinal products)?	○ Yes	◉ No	
3. Human cells / tissues			**Page**
Does this activity involve the use of human cells or tissues (not covered by section 1)?	○ Yes	◉ No	
4. Personal data			**Page**
Does this activity involve processing of personal data?	○ Yes	◉ No	
Does this activity involve further processing of previously collected personal data (including use of preexisting data sets or sources, merging existing data sets)?	○ Yes	◉ No	
Is it planned to export personal data from the EU to non-EU countries?	○ Yes	◉ No	
Is it planned to import personal data from non-EU countries into the EU or from a non-EU country to another non-EU country?	○ Yes	◉ No	
Does this activity involve the processing of personal data related to criminal convictions or offences?	○ Yes	◉ No	

4 - Other questions

Ethics Issues Table

5. Animals			Page
Does this activity involve animals?	○ Yes	◉ No	

6. Non-EU countries			Page
Will some of the activities be carried out in non-EU countries?	○ Yes	◉ No	
In case non-EU countries are involved, do the activities undertaken in these countries raise potential ethics issues?	○ Yes	◉ No	
It is planned to use local resources (e.g. animal and/or human tissue samples, genetic material, live animals, human remains, materials of historical value, endangered fauna or flora samples, etc.)?	○ Yes	◉ No	
Does this activity involve low and/or lower middle income countries, (if yes, detail the benefit-sharing actions planned in the self-assessment)?	○ Yes	◉ No	
Could the situation in the country put the individuals taking part in the activity at risk?	○ Yes	◉ No	

7. Environment, health and safety			Page
Does this activity involve the use of substances or processes that may cause harm to the environment, to animals or plants (during the implementation of the activity or further to the use of the results, as a possible impact)?	○ Yes	◉ No	
Does this activity deal with endangered fauna and/or flora / protected areas?	○ Yes	◉ No	
Does this activity involve the use of substances or processes that may cause harm to humans, including those performing the activity (during the implementation of the activity or further to the use of the results, as a possible impact)?	○ Yes	◉ No	

4 - Other questions

Ethics Issues Table

8. Artificial intelligence			Page
Does this activity involve the development, deployment and/or use of Artificial Intelligence-based systems? *if yes, detail in the self-assessment whether that could raise ethical concerns related to human rights and values and detail how this will be addressed.*	◉ Yes	○ No	24

9. Other ethics issues			Page
Are there any other ethics issues that should be taken into consideration?	○ Yes	◉ No	

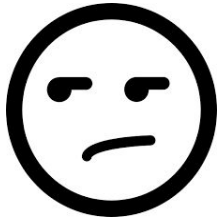
I confirm that I have taken into account all ethics issues above and that, if any ethics issues apply, I will complete the ethics self-assessment as described in the guidelines How to Complete your Ethics Self-Assessment ☒

All research activities must respect: □ the principle of proportionality □ the right to privacy □ the right to the protection of personal data □ the right to physical and mental integrity of all persons □ the right to equality and non-discrimination □ high levels of protection of the environment and human health.

Ethical dimension of the objectives, methodology and likely impact
N/A

Remaining characters 4997

Compliance with ethical principles and relevant legislation
There are no specific ethics issues that have a likely impact on project implementation. If any ethics issues apply, we will complete the ethics self-assessment as it is required by the regulations.



Remaining characters 4801

Ethics issues

Generally, ethics issues arise whenever research involves:

- humans participants, and/or their cells/tissues; personal data; animals; methods, materials or experiments that could harm the environment, research staff or participants;
- or when research is conducted **outside the EU**, especially in countries that lack adequate regulation and/or have a limited capacity to enforce the relevant ethical standards and guidelines.
- Ethics issues can also take the form of concerns about the **potential misuse** of new technologies, innovations, applications or research findings – even where research projects have benign intentions.

Serious & Complex Ethical Issues

Generally, ethics issues raised by research activities may be considered as "**serious**" when the proposed research, method(s), or outcome(s): □

- have the potential to **violate fundamental rights or freedoms** set out in the EU Charter of Fundamental Rights and European Convention on Human Rights, or undermine fundamental EU values such as human dignity, freedom, democracy, equality and the rule of law; or □
- have the potential to result in **significant harm to researchers, research participants, the public, animals or the environment**; or □
- in light of the European Code of Conduct for Research Integrity, fundamentally **call into question the integrity of the data and information included in the proposal or the integrity of the practices of the researchers**.

Serious & Complex Ethical Issues

Ethics issues raised by research activities may be considered as "**complex**" when the proposed research, method(s) or outcome(s): □

- involve the development or application of **particularly complicated methods or technologies that have not been sufficiently tested** and give rise to uncertainty as regards to the safety of participants and/or the impact of the expected results or outcomes on fundamental rights or research integrity; or □
- raise significant **ethics issues 'at scale'** – for example, due to the number of research participants, controversial methods, high-risk technologies or jurisdictions involved; or □
- raise multiple or '**intersectional**' ethics issues – meaning that the ethics issues may compound one another to exacerbate the potential impact on a particular group (e.g. research into marginalised or vulnerable groups that exposes them to the risk of stigmatisation, exclusion, reprisals or increased marginalisation).

ETHICS SELF-ASSESSMENT

8. Does this activity involve the development, deployment and/or use of Artificial Intelligence-based systems?

- Do project activities involve AI?
  - "'AI system' means 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 the input it receives, how to generate outputs such as **predictions, content, recommendations, or decisions that can influence physical or virtual environments**".
- Do we develop, deploy and/or use AI?

The AI ACT DOES (NOT) APPLY TO

Art. 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 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.

In any event, any research and development activity should be carried out in accordance with **recognised ethical and professional standards** for scientific research and should be conducted in accordance with applicable Union law.

See also recitals 25 and 97 (prototyping of general-purpose AI model).

Article 27 Fundamental rights impact assessment for high-risk AI systems

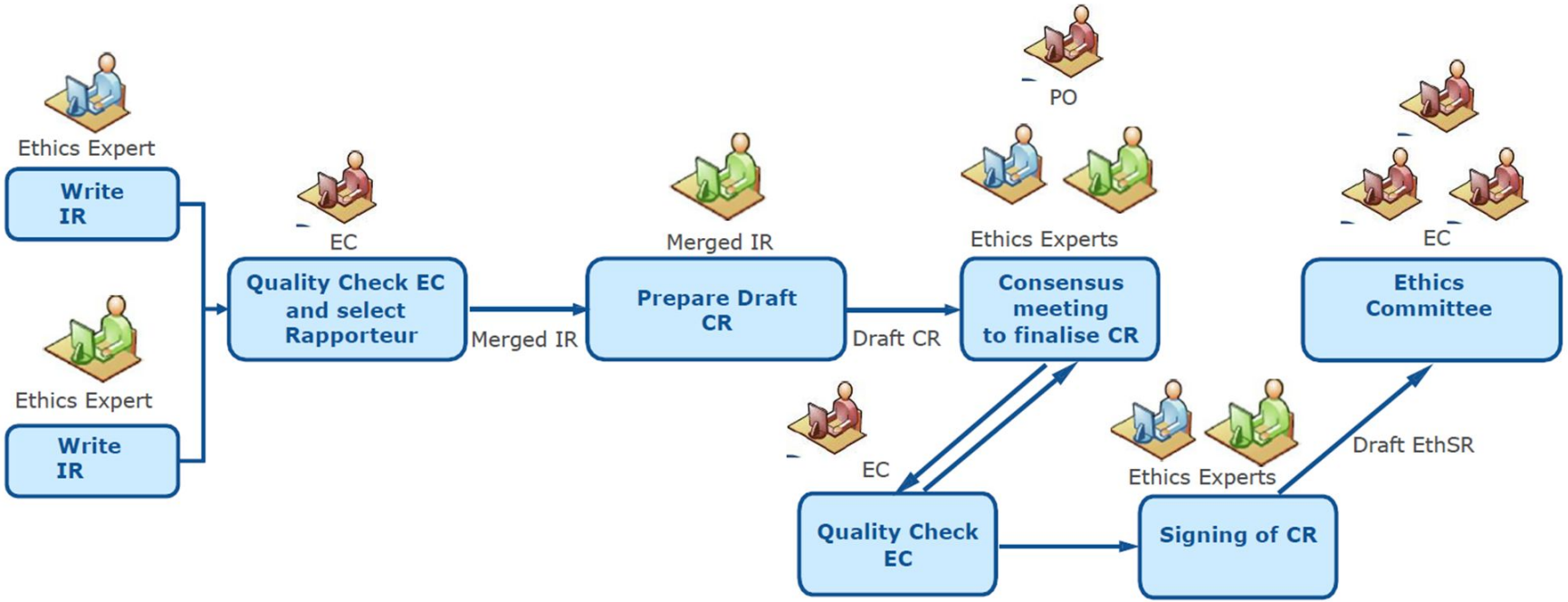
- 1. **Prior to deploying a high-risk AI system** referred to in Article 6(2) into use, with the exception of high-risk AI systems intended to be used in the area listed in point 2 of Annex III, deployers that are bodies governed by public law, or are private entities providing public services, and deployers high-risk AI systems referred to in points 5 (b) and (c) of Annex III, **shall perform an assessment of the impact on fundamental rights** that the use of such system may produce.

High risk AI ⇒ Risk Management

Article 6 - Classification rules for high-risk AI systems. AI system is intended to be used as a safety component of a product, or the AI system is itself a product. + Annex I (conformity assessment).

Article 9 - Risk management system. 1. A risk management system shall be established, implemented, documented and maintained in relation to high-risk AI systems.

Workflow of ethics screening



Ethics screening

- **Ethics clearance** —proposals without serious or complex ethics issues, and without need to update before grant signature.
- **Conditional ethics clearance** —proposals without serious or complex ethics issues but need for updates before grant signature. In the ethics Consensus Report, experts can *remind* the participants of the declarations they made and point out the standards and rules they are expected to comply with.
- **No ethics clearance (= go to ethics assessment)** —proposals that raise serious or complex ethics issues.

- The administrative effort must be in proportion to the risk involved, both for the Commission and for the beneficiaries of the grant.
- The choice between Ethics Check and Ethics Review should reflect the size of the grant and the importance of the ethics issues.
- The timing and number of Checks/Reviews depends on the structure, complexity and length of the project.

Ethics Advisor / Board / Mentor

- An "**Ethics Advisor**" (EA) is an **individual ethics expert** giving advice on issues of ethical gravity that relate to the planned and/or ongoing research in the context of an EU-funded project, and, if required, report to the Commission/Executive Agency/Funding Body.

- An "**Ethics Advisory Board**" (EAB) consists of three or more ethics experts, working together as panel in performing these tasks.

- Similar to the EA/EAB, an "**Ethics Mentor**" can be appointed to provide ethics guidance and advice. An Ethics Mentor can be a (senior) colleague, member of the same department or institution.

- Hence, **the key difference with an EA/EAB is that the Mentor may not be independent from the beneficiary and generally does not have any reporting duties**, although it is highly recommended that a report on the activities of the Mentor is kept on file.

Ethics Advisor / Board / Mentor

- The mandate of the EA/EAB should be clearly defined and outlined in a Memorandum of Understanding (MOU).
- For EAs/EABs required by the Ethics Review, the **EA and EAB Members are not and may not be held responsible for the Beneficiary/ies' compliance** with the ethics requirements and the applicable ethical and legal standards.
- As EAs/EAB Members **cannot** work so closely with the Beneficiary/ies or **become "part" of the working team**, it might be worthwhile considering the use of individual work packages and integration of ethics experts as project partners as part of the working structure that ensures easier integration of the expertise in the daily research.

Unlawful activities are not funded

REMEMBER that actions for the development of products and technologies whose use, development or production is **prohibited by applicable international law** cannot be eligible for funding.

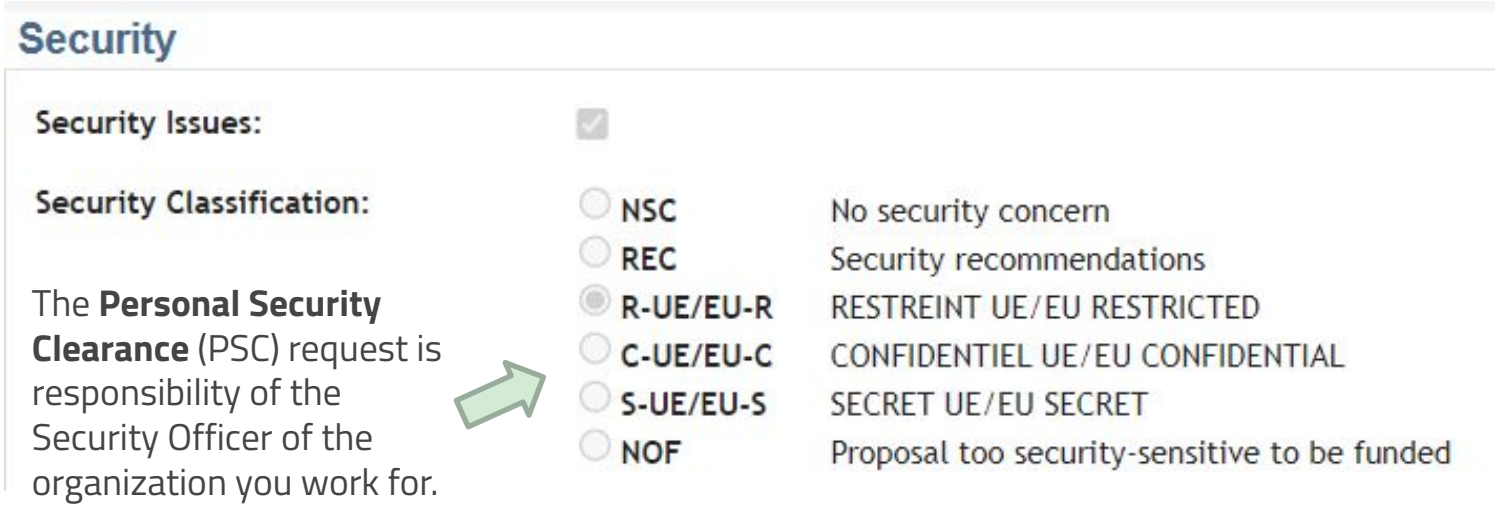
Lethal autonomous weapons without the possibility for meaningful human control over engagement decisions cannot be funded under EDF calls. HORIZON has an exclusive focus on civil application. Dual-use technologies must be dealt accordingly.

AI LAW: "art. 3. This Regulation **does not apply** to AI systems where and in so far they are placed on the market, put into service, or used with or without modification exclusively for **military, defence or national security purposes**, regardless of the type of entity carrying out those activities."

Grant Management System / Ethics

You must ensure respect for people and for human dignity and fair distribution of the benefits and burden of research, and that you must protect the values, rights and interests of the research participants.

Grant Management System / Security



The **Personal Security Clearance** (PSC) request is responsibility of the Security Officer of the organization you work for.

CNI Digital Office

For projects with EU classified information (EUCI) that require security clearances, it should be made clear that the members of the **EAB or the EA might also need adequate clearance** in order to have access to all relevant information.

Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment

- The Ethics Guidelines introduced the concept of Trustworthy AI, based on seven key requirements:
  1. Human agency and oversight
  2. Technical robustness and safety
  3. Privacy and data governance
  4. Transparency
  5. Diversity, non-discrimination and fairness
  6. Environmental and societal well-being and
  7. Accountability.
- Through the Assessment List for Trustworthy AI (ALTAI), AI principles should be translated into an accessible and dynamic checklist that guides developers and deployers of AI in implementing such principles in practice.

AI/Information Ethics

- **Beneficence** means beneficial to humanity including the protection of the planet on which humanity thrives.
- **Non-maleficence** is invoked in injunctions against infringements of privacy and other irresponsible uses of AI, which may apply to "accidental" and "deliberate" harms.
- **Autonomy** refers to the power to *decide which decisions to take.*
- **Justice** implies measures to prevent risks of "unfair discrimination", unfair distribution of benefits, and other cumulative negative outcomes.
- **Explicability** is the crucial missing piece of the AI ethics jigsaw.
  - Explicability is understood as incorporating both the *epistemological* sense of ***intelligibility***—as an answer to the question 'how does it work?'—and in the *ethical* sense of ***accountability***—as an answer to the question 'who is responsible for the way it works?'.
  - The principle of 'explicability' incorporates both the **epistemological** sense of 'intelligibility' and the **ethical** sense of 'accountability'.

Floridi, L., 2023. The Ethics of Artificial Intelligence: principles, challenges, and opportunities. Oxford University Press.

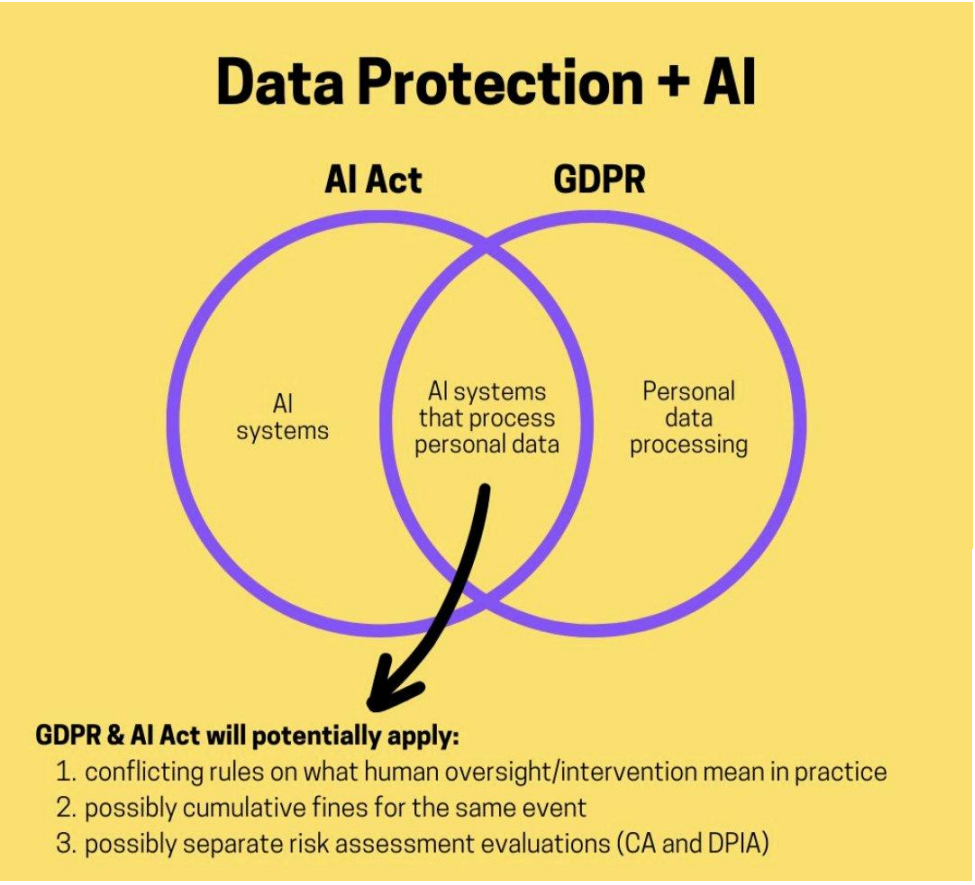
Resources

- Ethics By Design and Ethics of Use Approaches for Artificial Intelligence
- SHERPA - Guidelines for the Ethical Development of AI and Big Data Systems: An Ethics by Design approach
- Data Protection Decision Tree
- Ethics in Social Science and Humanities
- Research Ethics in Ethnography/Anthropology

AI-HLEG 2019

Fundamental Rights of Human Beings	Ethical Principles five principles and correlated values that must be observed to ensure that AI is developed in a human-centric manner	Achieving Trustworthy AI means that the general and abstract principles need to be mapped into concrete requirements for AI systems and applications	Methods to ensure values-by-design
• Respect for human dignity • Freedom of the individual • Respect for democracy, justice and the rule of law • Equality, non-discrimination and solidarity including the rights of persons belonging to minorities • Citizens rights.	The Principle of: • Beneficence: "Do Good" • Non Maleficence: "Do no Harm" • Autonomy: "Preserve Human Agency" • Justice: "Be Fair" • Explicability: "Operate transparently".	1. Accountability 2. Data Governance 3. Design for all 4. Governance of AI Autonomy (Human oversight) 5. Non-Discrimination 6. Respect for (& Enhancement of) Human Autonomy 7. Respect for Privacy 8. Robustness 9. Safety 10. Transparency.	Examples of values-by-design are Privacy-by-design or Security-by-design). The requirements for Trustworthy AI need to be "translated" into procedures and/or constraints on procedures, which should be anchored in an intelligent system's architecture. *Non-Technical Methods* (Regulation; Standardization; Accountability Governance; Codes of Conduct; Education and awareness to foster an ethical mind-set; Stakeholder and social dialogue; Diversity and inclusive design teams).

CHARTER OF FUNDAMENTAL RIGHTS OF THE EUROPEAN UNION (2010/C 83/02).



Art. 9 GDPR
Processing of special categories of personal data

Personal data revealing racial or **ethnic** origin, **political** opinions, **religious** or **philosophical** beliefs, or **trade union** membership, and the processing of **genetic** data, **biometric** data for the purpose of uniquely identifying a natural person, data concerning **health** or data concerning a natural person's **sex life** or sexual orientation.

'personal data' means **any information relating to an identified or identifiable natural person** ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

Example of a special category of data

Data Privacy Manager

BIOMETRIC DATA

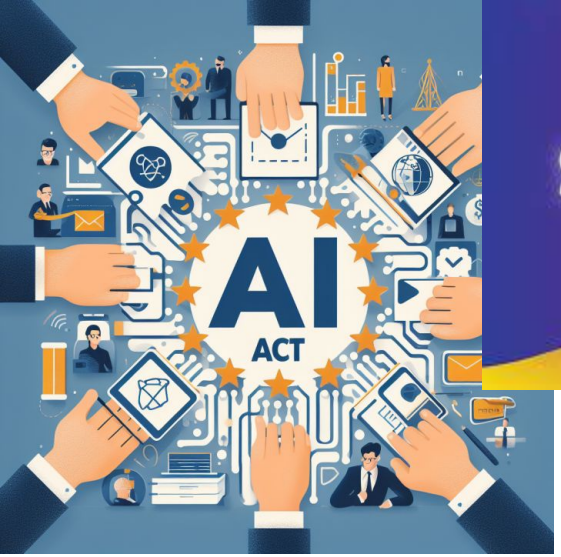
- facial recognition
- fingerprints
- voice recognition
- iris scanning
- palmprint verification
- retina recognition
- ear shape recognition

HEALTH DATA

- patient medical history
- data on disability
- illnesses,
- medical diagnosis,
- medical treatment,
- medical opinions
- fitness tracker data

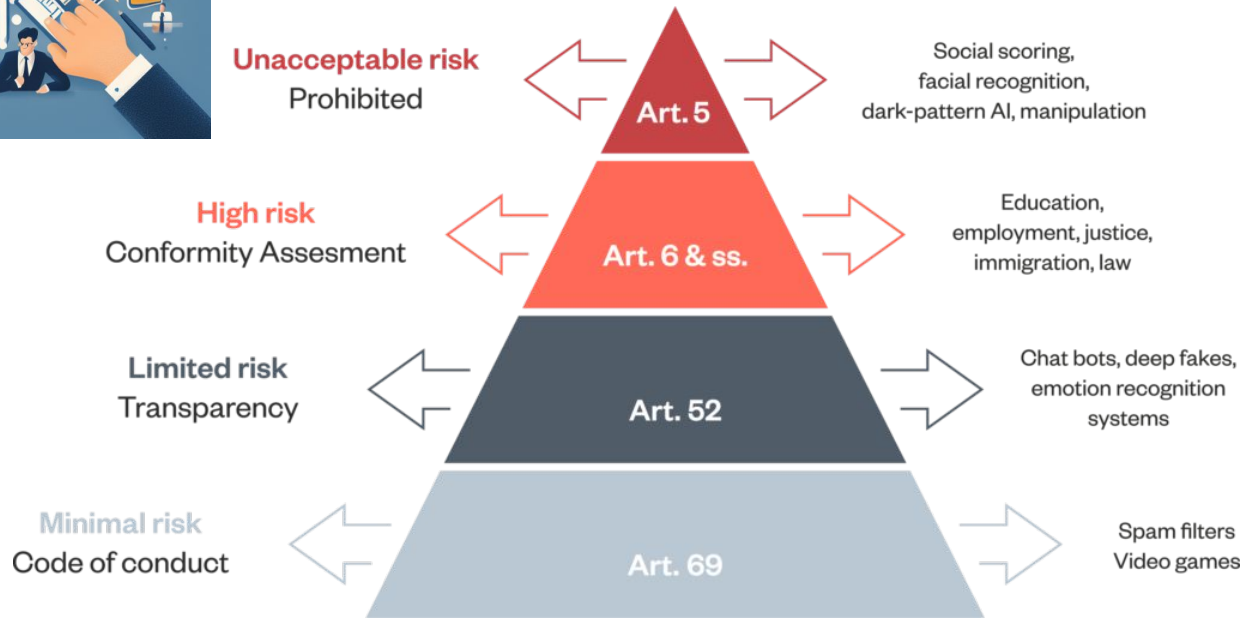
GENETIC DATA

- chromosomal analysis
- DNA analysis
- RNA analysis


AI
ACT


DIGITAL
SERVICES
ACT
#DigitalEU


GDPR
Data Protection Officer (DPO)
Compliance
25 May 2018
Data Breaches
Personal Data


Unacceptable risk
Prohibited
Art. 5
Social scoring,
facial recognition,
dark-pattern AI, manipulation
High risk
Conformity Assesment
Art. 6 & ss.
Education,
employment, justice,
immigration, law
Limited risk
Transparency
Art. 52
Chat bots, deep fakes,
emotion recognition
systems
Minimal risk
Code of conduct
Art. 69
Spam filters
Video games

Article 5 – Prohibited AI Practices

1. The following AI practices shall be prohibited:

- (a) … subliminal techniques beyond a person's consciousness **or purposefully manipulative or deceptive techniques**, … **impairing their ability to make an informed decision**, thereby causing … another **person or group of persons** significant harm;
- (b) **exploits any of the vulnerabilities** of a person or a specific group of persons due to their age, disability or a specific social or economic situation, … materially distorting the behaviour of that person .. *significant harm;*
- (c) evaluation or classification of *natural persons or groups of persons … inferred* or predicted personal or personality characteristics … detrimental or unfavourable treatment;
- (d) likelihood of a natural person committing a criminal offence, based solely on the **profiling** of a natural person or on assessing their personality traits and

Article 5 - Prohibited AI Practices

1. The following AI practices shall be prohibited:

- (e) AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage;

- AI systems to **infer emotions** of a natural person in the areas of **workplace and education institutions**, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons.

- **biometric** categorisation systems … 'real-time' remote biometric identification systems in publicly accessible spaces … unless the targeted search for specific victims of abduction, trafficking in human beings or sexual exploitation of human beings, as well as searching for missing persons;

ANNEX III – High-risk AI systems

AI systems intended to be used:

- for **biometric verification and categorisation** (see GDPR art. 9 Processing of special categories of personal data) or **emotion recognition**;
- as **safety components** in the management and operation of **critical digital infrastructure**, road traffic, or in the supply of water, gas, heating or electricity;
- for **education and vocational training,** students' admission, evaluation of **learning outcomes**, assessment of students' performance, or **monitoring** and **detecting prohibited behaviour of students;**
- for **employment, workers management and access to self-employment**, for the **evaluation or termination of job candidates,** the **filtering of job applications..**

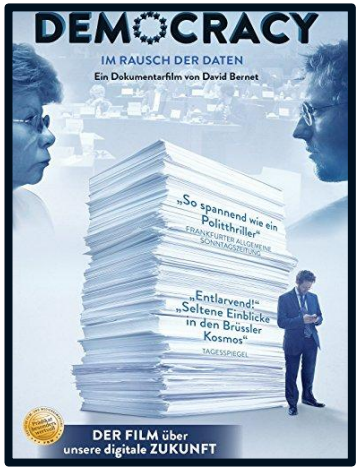
ANNEX III - High-risk AI systems

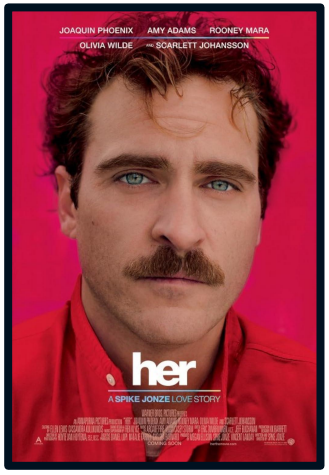
- AI systems intended to be used for access to and enjoyment of **essential private services** and **essential public services and benefits:**
  - for assessing the **eligibility** of people to **essential public services,** for assessing **creditworthiness** or detecting financial fraud, for risk assessment and pricing of **life and health insurance,** for the evaluation and classification of **emergency calls,** assess a natural **person's risk of becoming the victim of criminal offences.**
- AI systems intended to be used for **law enforcement:**
  - to evaluate the reliability of evidence, the risk of becoming **victim of criminal offences**, likelihood of a natural person of offending or reoffending, **suspects' profiling**, any kind of **polygraphs**.

ANNEX III - High-risk AI systems

- AI systems intended to be used for **migration, asylum and border control management:**
  - for assessing eligibility of the natural persons applying for a status, assessing if they represent a risk to public security or health, or for detecting, recognising or identifying natural persons.
- AI systems intended to be used for the **administration of justice and democratic processes:**
  - to assist a **judicial** authority **in researching and interpreting facts and the law** and in applying the law to a concrete set of facts, or to influence the **outcome of an election** or referendum.


Science Fiction


DEMOCRACY
IM RAUSCH DER DATEN
Ein Dokumentarfilm von David Bernet
„So spannend wie ein Politthriller"
FRANKFURTER ALLGEMEINE SONNTAGSZEITUNG
„Entlarvend!" „Seltene Einblicke in den Brüssler Kosmos"
TAGESSPIEGEL
DER FILM über unsere digitale ZUKUNFT

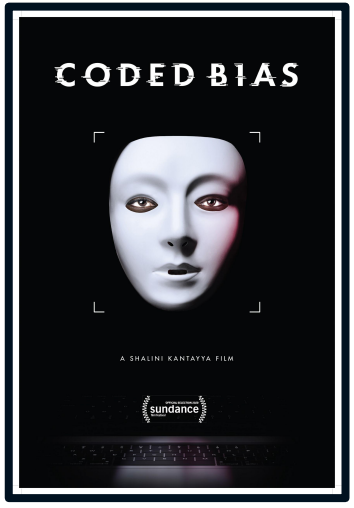

JOAQUIN PHOENIX AMY ADAMS ROONEY MARA
OLIVIA WILDE and SCARLETT JOHANSSON
her
A SPIKE JONZE LOVE STORY

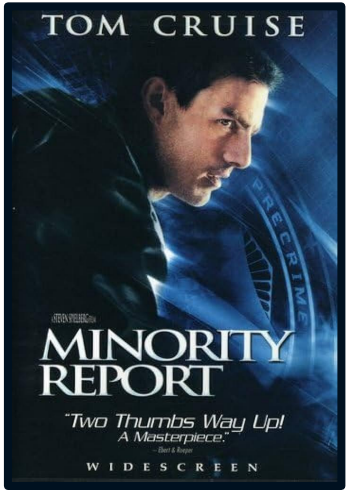

HARRISON FORD
BLADE RUNNER

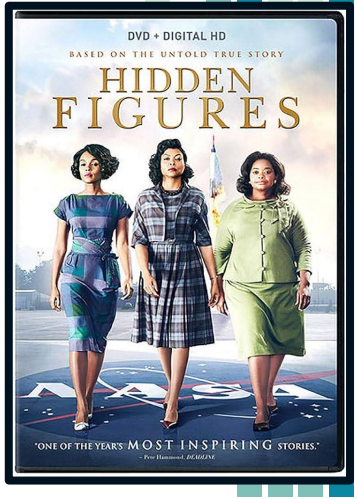

KEANU REEVES LAURENCE FISHBURNE
MATRIX
ON MARCH 31ST THE FIGHT FOR THE FUTURE BEGINS.


THEY TOOK YOUR DATA. THEN THEY TOOK CONTROL.
A NETFLIX ORIGINAL DOCUMENTARY
THE GREAT HACK
From the Academy Award® nominated filmmakers behind THE SQUARE
sundance
JULY 24 | NETFLIX


sundance
the dark side of social media...
from the people who created it.
/the social dilemma_
NETFLIX


CODED BIAS
A SHALINI KANTAYYA FILM
sundance


TOM CRUISE
MINORITY REPORT
"Two Thumbs Way Up! A Masterpiece."
WIDESCREEN


DVD + DIGITAL HD
BASED ON THE UNTOLD TRUE STORY
HIDDEN FIGURES
"ONE OF THE YEAR'S MOST INSPIRING STORIES."

THANKS!

Any question?

You can find us at:
sara.degli.esposti@csic.es
pompeu.casanovas@iiia.csic.es