



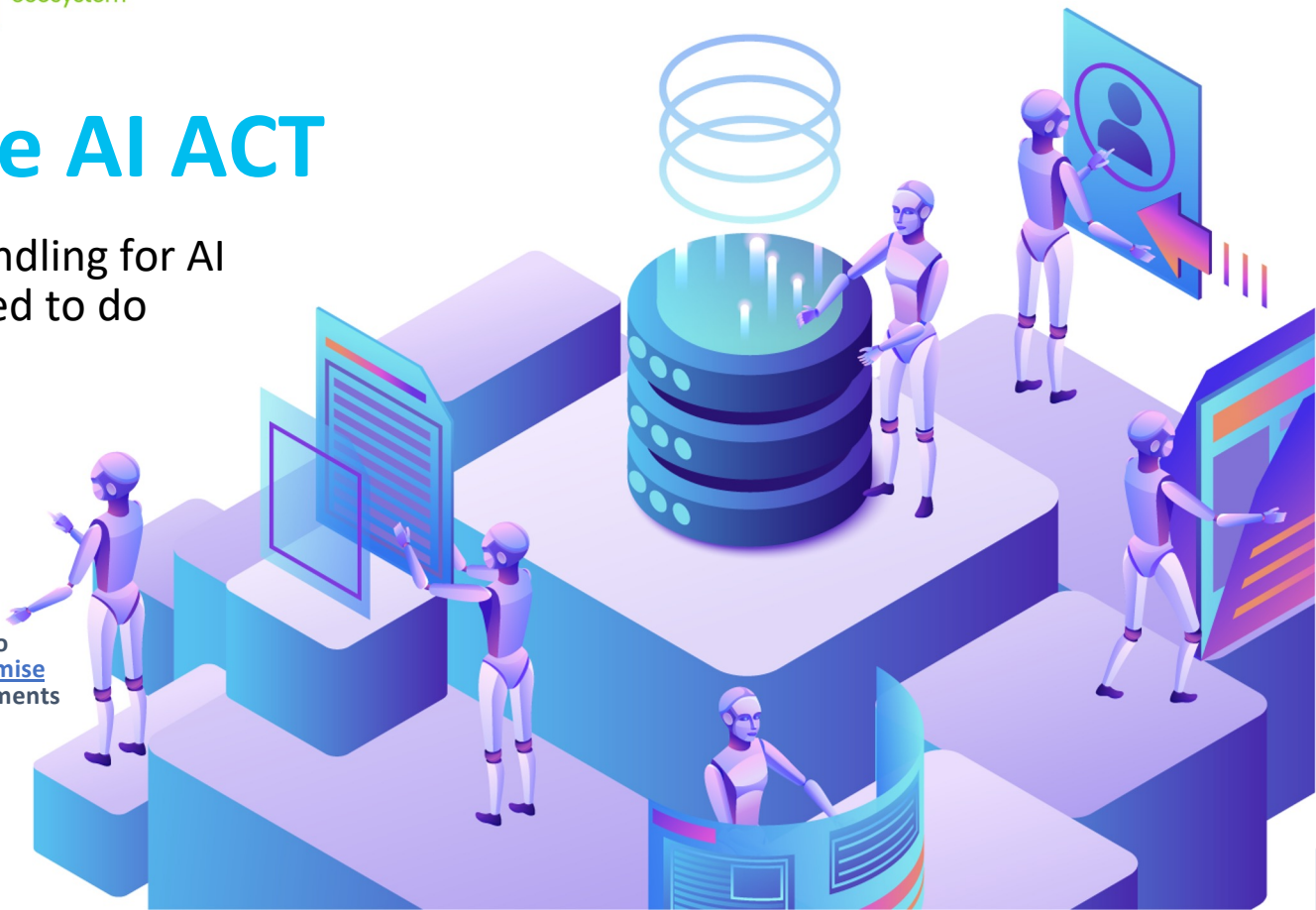
AI, Data and Robotics
ecosystem

Leveraging the AI ACT

Introduction to why risk-handling for AI is needed and what you need to do

Version 20230822 and Disclaimer:

We use partial quotations and also references to articles (e.g. **#A1** for article Nr. 1 of the [compromise draft AI ACT published 11 May 2023](#)), but statements are interpretations by the authors. You should ultimately read the AI Act when it is finalized in Q4 2023. See [EUP news](#) and [\[01\]](#) for updates.



Funded by
the European Union

Contents

/1/	Introduction
/2/	Terms
/3/	Using the AI Act: Judge risks; Comply with rules
/4/	AI Act: Life-Cycle for (high-risk) AI Solutions
/5/	AI Act: Duties of Providers
/6/	AI Act: Duties of Deployers
/7/	Deploying AI with Foundation or Large Language Models
/8/	Ensuring appropriate Risk Management System
/9/	Ensuring appropriate Data Governance
/10/	Ensuring Transparency
/11/	Enabling Human oversight
/12/	Enabling Accountability: Documentation & Record-keeping
/13/	Additional Issues
/14/	Leveraging the AI Act
/15/	References & Further reading

ABSTRACT: *This presentation gives an overview and further reading for the EU AI Act – drafted in February 2021 [02], has been extensively discussed in the EU Council and Parliament [01], is due for final agreement in late 2023, and will be enforced by the end of 2025.*

The AI Act is based on broad horizontal classifications of risks, focusing on steps to mitigate the highest risks and forbid some use cases. The goal is to protect human rights, public safety, and environment, from misuse, unreliable operation, or bias of an AI system.

Reading the AI ACT, the preliminaries explain the legal basis, scope, and concepts, then there are sections (Articles) with specific requirements on risk management, data quality, oversight, operational monitoring, documentation, etc.

Requirements are most strict for high-risk scenarios, which are broadly defined but subject to case-by-case review by national Notified Bodies before deployment.

AI producers and deployers are subject to significant fines for breaking the rules and thereby harming EU citizens, no matter where in the world they are (similar approach to GDPR) . Harmonised standards are under development to operationalize the rules, for product conformance.

/1/ Introduction

What is AI?

“AI system means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments.” #A3(1)

Why the AI ACT is needed

- Protect citizen fundamental rights, health and safety, the environment. Encourage innovation INSIDE the rules.
- Horizontal Legislation is designed to make “level playing field” across all industries, avoid “case-by-case” law
- Allow development of harmonised EU Norms so businesses will know exactly what is needed for “presumption of conformity” #A40, #A42 for free movement of AI-based goods and services cross-border in the EU

Who is impacted, who needs to act

- Protect all citizens of EU, anywhere in the world, from AI
- All EU/EEA organisations making AI must follow the rules (risk management, monitoring, transparency, etc.)
- Non-EU businesses offering services to EU citizens (anywhere) must ALSO follow same rules!
- Producers of AI software (B2B sales) and Deployers (B2C sales) have slightly different duties
- SMEs will need AI to stay competitive, but they have less resources for compliance → EU help for SMEs #A55






Reading this presentation ... K.I.S.S. principle

- **EMPHASIS HERE is on HIGH-RISK Solutions (medium- or low-risk has fewer mandatory requirements)**
- We keep it short and simple! ... but we give references [03] to further reading. Many terms are NOT yet fully clarified.
- Everything here is based on the EU Parliament approved compromise from 11 May 2023 available [here](#) [04]

**TIMING: AI Act drafted 02/2021, analysed by EU Council, reformulated by EU Parliament, 07/2023 enter Trialogues
Expect final agreement in Q4 2023, Legal enforcement in Q4-2025**



/2/ Terms

-  **'user'** means any natural or legal person, including a public authority, agency or other body, under whose authority the system is used;
-  **'provider'** places on the market or puts into service an AI systems in the Union, irrespective of whether those providers are established within the Union or in a third country; **#A2(a), #A3(2)**
-  **'deployer'** means any natural or legal person, public authority, agency or other body using an AI system under its authority, except where the AI system is used in the course of a personal non-professional activity; **#A3(4)** (**was termed as 'user' in early versions*)
- 'distributor'** means any natural or legal person in the supply chain, other than the provider or the importer, that makes an AI system available on the Union market without affecting its properties; **#A3(7)**
-  **'authorised representative'** means any natural or legal person established in the Union who has a written mandate from a provider of an AI system to, respectively, perform and carry out on its behalf the obligations and procedures established by this Regulation; **#A3(5)**
- 'importer'** means any natural or legal person physically present or established in the Union that places on the market an AI system that bears the name or trademark of a natural or legal person established outside the Union; **#A3(6)**
- 'operator'** means the provider, the deployer, the authorised representative, the importer and the distributor;
- 'product manufacturer'** means a manufacturer within the meaning of any of the Union harmonisation legislation listed in Annex II;
- 'notifying authority'** means the national authority responsible for setting up and carrying out the necessary procedures for the assessment, designation and notification of conformity assessment bodies and for their monitoring; **#A3(19)**
-  **'market surveillance authority'** means the national authority carrying out the activities and taking the measures pursuant to Regulation (EU) 2019/1020; **#A3(26)**

Note that in USA and other countries similar terms are used in different ways. See e.g. [\[05\]](#), [\[06\]](#)



/2/ Terms (continued)

‘**risk**’ means the combination of the probability of an occurrence of harm and the severity of that harm; [#A3\(1a\)](#)

‘**substantial modification**’ means a change to or a series of modifications of the AI system after its placing on the market or putting into service which is not foreseen or planned in the initial risk assessment by the provider and as a result of which the compliance of the AI system with the requirements set out in Title III, Chapter 2 of this Regulation is affected or results in a modification to the intended purpose for which the AI system has been assessed; [#A3\(23\)](#)

‘**post-market monitoring**’ means all activities carried out by providers of AI systems to proactively collect and review experience gained from the use of AI systems they place on the market or put into service for the purpose of identifying any need to immediately apply any necessary corrective or preventive actions; [#A3\(25\)](#)

‘**foundation model**’ means an AI system model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks; [#A3\(1c\)](#)

‘**general purpose AI system**’ means an AI system that can be used in and adapted to a wide range of applications for which it was not intentionally and specifically designed; [#A3\(1d\)](#)

‘**biometric identification**’ means the automated recognition of physical, physiological, behavioural, and psychological human features for the purpose of establishing an individual’s identity by comparing biometric data of that individual to stored biometric data of individuals in a database (one-to-many identification); [#A3\(33b\)](#)

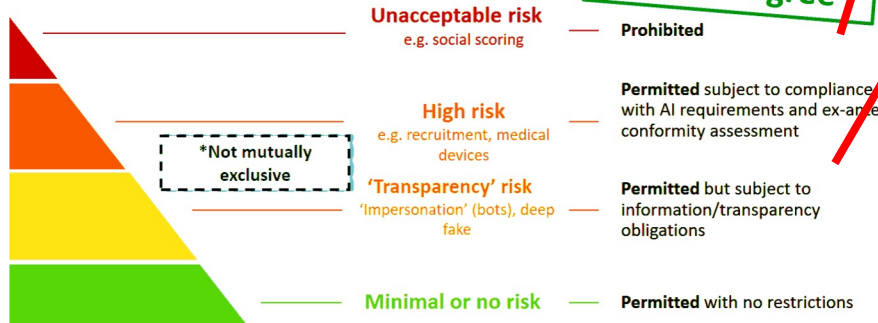
‘**trilogue**’ is an informal interinstitutional negotiation between the European Parliament, the Council of the European Union and the European Commission, to compose a legislative proposal, for adoption by each of those institutions’ formal procedures; [\[07\]](#)

/3/ Using the AI Act: Judge risks; Comply with rules

Classification of AI system based on the level of risk

- **Unacceptable risk:** Prohibited (see list xxx)
- **High risk:** Permitted, subject to compliance/obligations
 - MUST be certified by a third-party (Notified body)
- **Low risk:** Permitted, with no restrictions
 - Deployer is advised to perform a risk assessment

Risk-based + horizontal approach



Prohibited: remote biometric identification in publicly-accessible spaces for law enforcement without human intervention (exceptions exist); subliminal or purposefully manipulation of people causing significant harm; social scoring and predictive policing; face recognition based on indiscriminate web-scraping or CCTV footages **#A5(1)**

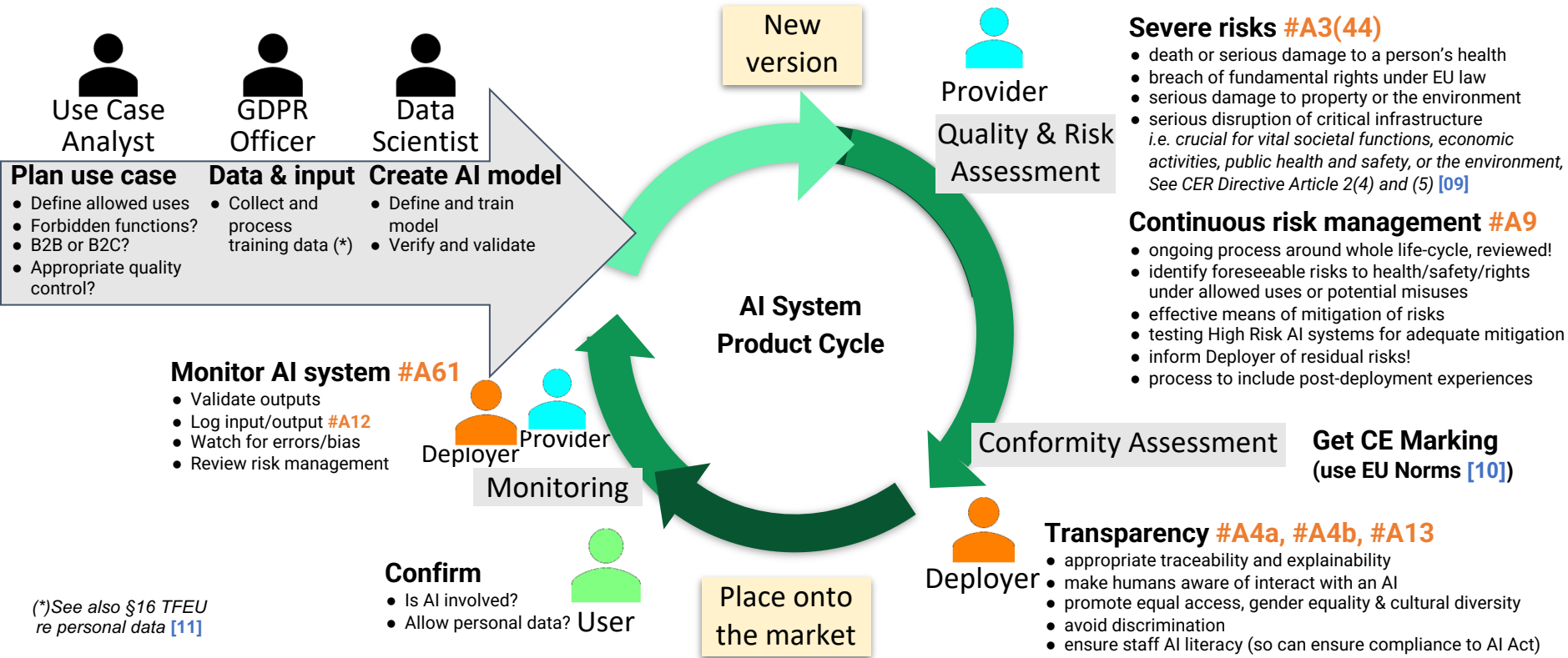
High Risk: health/life insurance; digital infrastructure; emotion recognition (if not prohibited); student monitoring; national border management; prediction of border crossings; elections; recommender systems in very large social media platforms; General purpose AI & foundation models #A28b will get tougher constraints.

See scientific literature for further analysis **[08]**

Exceptions: permit some high-risk areas if research, open-source, pre-deployment sandboxes.

Figure source page 5 of <https://www.ihk.de/blueprint/servlet/resource/blob/5198826/5e5feb7229bcdfba556b26cae8f95ecf/p01-eu-com-salvatore-scalzo-data.pdf>

/4/ AI Act: Life-Cycle for (high-risk) AI Solutions #A3(1a)



(*)See also §16 TFEU re personal data [11]



/5/ AI Act: Duties of Providers

Provider

1. **Ensure your intended AI system does not involve prohibited data or features #A5**
2. **Check if your AI system is High Risk** for intended use cases (**#A6** and **Annex III**)
 - 1) risk of harm to the health or safety of EU citizens
 - 2) risk of adverse impact on fundamental rights
1. **If “Yes, High Risk”: ensure comply to requirements for High Risk AI! (Chapter 2 of Title III)**
 - 1) Check Obligations (Title III, Chapter 3) and then document compliance with requirements (Title III, Chapter 2);
 - 2) Establish a Risk Management System
 - 3) Perform risk assessment (= probability x values at risk)
 - 4) Ensure risk mitigation and monitoring based on processes using
 - Transparency, Human Oversight, Accuracy, Robustness and Cybersecurity
 - 5) Prepare Technical Documentation for Deployer (see page 15 below!)
 - Scope of (allowed) use cases, limitations of the AI Solution, means of monitoring
 - 1) Collaborate with Deployer for post-deployment monitoring



/6/ AI Act: Duties of Deployers

Deployer

Are you a Deployer of AI systems?

- 1) Ensure all “Duties of Providers” have been documented for you

If “Yes, High Risk”: ensure

- 1) Transparency rules satisfied in your use case **#A4a, #A4b**
- 2) Your processes include appropriate traceability and explainability
- 3) The users (humans) are made aware if they interact with an AI
- 4) Check equal access, gender equality & cultural diversity respected
- 5) Avoid discrimination (and check for biases)
- 6) Ensure staff AI literacy enough to ensure compliance to AI Act
- 7) If human oversight is required, ensure it is timely and logged
- 8) All info needed in **#A51(1a)(a), #A51(1b), #A28b(e)** are provided and kept updated



/7/ Deploying AI with LLM or Foundation Models (high-risk by definition)

Provider



Deployer

Special regulations for Large Language Model and Foundation Model are still under discussion. See Recital 60 g, Recital 60 h

What is mandatory?

- Disclosing that content was generated by AI
- Ensuring content is labelled as coming from AI (e.g. watermark)
- Designing the model to prevent it from generating illegal content
- Monitoring that copyright/personal data used in training is done with permission
- Publishing summaries of copyrighted data used for training

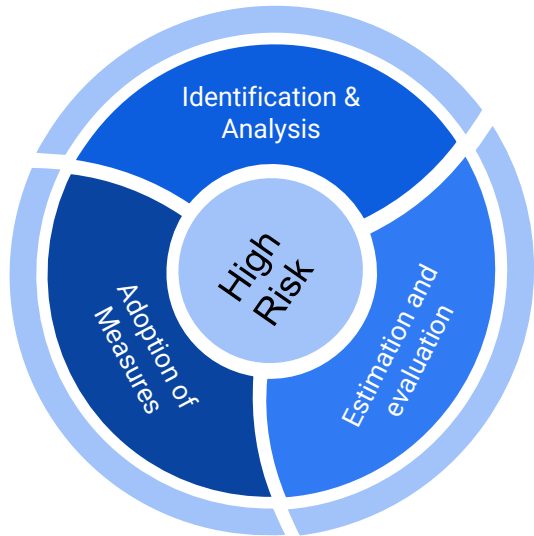
Relevant Roles

- Foundation Model provider **#A28b**
- 3rd-party AI component/service supplier **#A28(3)**
- AI integrators for high risk under **#A6(2)** need to comply with **#A16**
- AI deployer **#A29**
- Legal person affected by AI system **#A28a #A28b #A28c**

/8/ Ensuring appropriate Risk Management System for high-risk AI Systems #A9

See also [12]

Risk Management = Continuous iterative process run throughout the entire lifecycle to ensure its continuing effectiveness, and documentation of any significant decisions and actions taken



How (Measures to be ensured)?

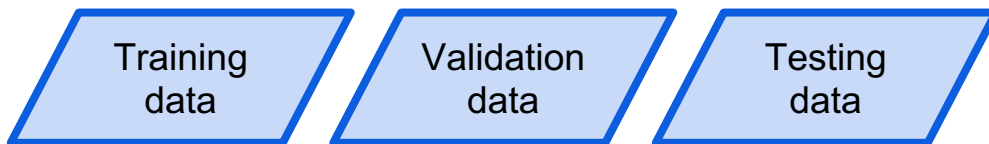
- **Identify** and analyse known and foreseeable risks, **also under misuse**
- **Analyse** risks from post-market monitoring **#A61**
- **Eliminate** risks as far as possible through adequate design and development
- **Mitigate unavoidable risks** with control measures in relation to degree of risk
- **Provide adequate information** (see **#A13**) + appropriate **training to users**
- **Document** all significant decisions and relevant data
- **Testing prior to deployment !**

Additional Considerations:

- Estimate and ensure relevant residual risk associated with each hazard as well as the overall residual risk of the system is reasonably judged to be acceptable
- Specifically consider whether the high-risk AI system is likely to adversely **impact vulnerable groups of people or children**
- Carry out a **fundamental rights impact assessment #A29a**
- **Report** any serious incident to appropriate authorities **#A62, #A65(1)**

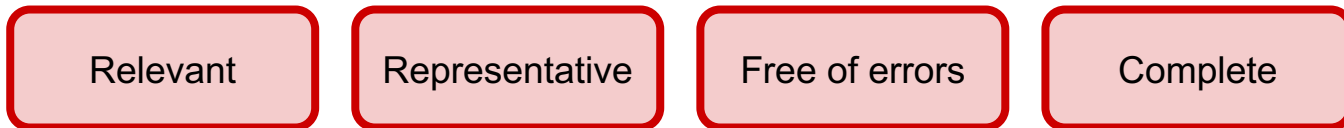
/9/ Ensuring appropriate Data Governance #A10

During their life cycle (acquisition, processing and use, sharing, disposal), effectively manage:

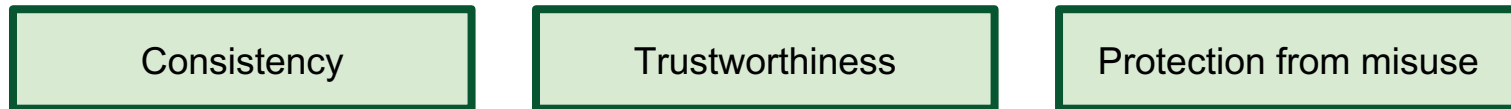


* If contains special categories of personal data, must have appropriate fundamental rights safeguards
 * If contains copyrighted work, must publicly provide summaries #A28b(4)(c)

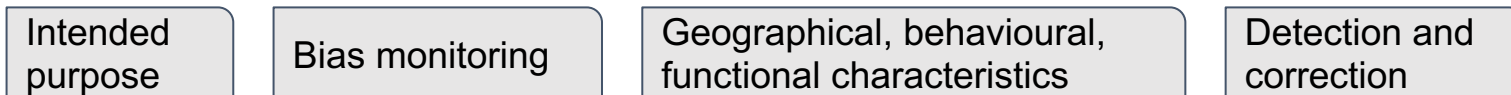
To meet quality criteria:



To ensure:



Considering:





/10/ Ensuring Transparency #A13, #A52

See also [01], [13]

W
H
E
N

Transparency Obligations apply
for high-risk systems that

- interact with humans,
- detect emotions,
- determine (social) categories based on biometric data
- generate/manipulate content ('deep fakes')

W
H
A
T

and require that

- allow persons to ***make informed choices*** or ***step back*** from a given situation
- people must be informed that the content is generated.

Success is when high-risk AI systems

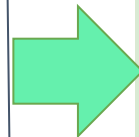
- enable users to interpret the system's output and use it appropriately; #A13(1)
- have/show instructions for use, such as
 - identity and contact details of the provider;
 - characteristics, capabilities and limitations of performance (intended purpose, level of accuracy, robustness and cybersecurity, risks to the health and safety or fundamental rights, performance, specifications for the input data);
 - changes/updates to the high-risk AI system;
 - human oversight measures
 - expected lifetime, maintenance and care measures.

/11/ Enabling Human Oversight #A14

Purpose: Ensure that high-risk AI systems are subject to appropriate levels of human control and that humans can override or deactivate them if needed.

How ensured:

- Provider shall identify and build into the system, when technically feasible, measures for human intervention to prevent/minimize risks to health, safety, and fundamental rights
- Provider shall inform Deployer / User measures appropriate for them to implement.



NOTE: many many regulations already assume some oversight and multi-level maturity models.

Results

(1) Human in the loop

Enable the 'human overseer' to

- spot anomalies
- be aware of AI limitations
- correctly interpret the system's outputs
- override/disregard the system

(1) Human override as needed

- override/disregard the system

Especially, a key aim is to prevent or minimise risks to fundamental rights.

/12/ Enabling Accountability: Documentation and Record-keeping



Before system is placed on the market

While system is operating

Technical documentation

#A11, #A18

Title III, Chapter 2



Provider

Demonstrate #A8-15 compliance (req. for high-risk AI systems)

- Min. elements set out in **Annex IV**
- 1) general description (incl. intended purpose, version, instructions of use for Deployer #A13)
- 2) detailed desc. of system + of development process (incl. assessment of #A14, #A13(3)(d))
- 3) detailed info about monitoring, functioning, and control (incl. human oversight #A14)
- 4) detailed desc. of risk mgmt. system #A9
- 5) desc. of change made through its lifecycle
- 6) a list of standards applied in full or in part
- 7) a copy of the EU declaration of conformity
- 8) detailed desc. of system to evaluate performance in post-market #A61, incl. monitoring plan #A61(3)

Record-keeping

#A12, #A20



Provider

Enable automatic recording of events ('logs') for traceability throughout AI lifecycle



Provider

Enable the monitoring of risk as in #A65(1) and facilitating post-market monitoring as in #A61

Post-market monitoring

#A20(4), #A61, #A62



Provider

Evaluate continuous #A8-15 compliance



Provider

Actively and systematically collect, document, analyse relevant data for evaluation
 Deployer provide such data



Provider
Deployer

Monitor, report if #A65(1) risk or #A62 serious incident/malfunction
 Deployer inform Providers and stop using system



Provider

notify Market Surveillance Authorities, within 15 days after become aware



Provider

Establish monitoring system based on monitoring plan in **Annex IV** tech. doc.



/13/ Additional Issues

Some features are difficult to quantify and validate, depending on the use case

- Accuracy and Robustness and Cybersecurity #A15, #A56(2)(a)

Some administrative and regulatory process are very complex

- Using and regulating the “proof of concept” Sandboxes #A53, #A54
- Difficulties in regulating large language models and foundation frameworks, like GPT4, ChatGPT, LLaMA [14]
- Accreditation of Third-Party Test Centres
 - EU has > 6000 Notified Bodies [15] [16] for third-party conformity assessment (CE label).
 - How many test centres will handle AI ? How accredited?

Some use cases are very complex

- Metaverse® is a large-scale, non-EU based, use case
 - Privacy? Are my preferences and personal data used according to GDPR and permissions?
 - Transparency? Is that avatar from a real person or just an AI?
 - Recommendation system built-in?
 - Is an SME that has a point-of-sale in a Metaverse responsible for all the Metaverse features?



/14/ Leveraging the AI Act

The AI Act has costs, but many benefits ... leverage them!

Costs

- Much care needed before deployment for Risk Analysis and mitigation
- Needs careful documentation of risks, mitigations, processes, operations
→ follow EU Norms as soon as available [10]
- For high-risk use cases, need to pay for 3rd party conformance testing



Benefits

- **Society** sees far fewer solution failures, scandals and disappointments
→ biases/errors etc can cost lives (e.g. in medical diagnostics)
→ lower risk of a “counter reaction”, even banning of solutions
→ huge improvement in detection/mitigation of “AI for crime”
- **Providers and Deployers** get clear responsibilities
→ market is more transparent; liability is separated
- **End-users and businesses** see Certifications, gain trust
→ more system biases will be caught/fixed before deployment
→ deployment and expansion can be more rapid and certain



/15/ References & Further reading

- [01] Comparison of 02/2021 draft of the AI ACT with 11/05/2023 negotiated result of the EU Parliament and with 11/2022 deliberations of EU Council. See <https://www.europarl.europa.eu/cmsdata/272920/AI%20Mandates.pdf>
- [02] 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. COM/2021/206 final. See <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>
- [03] Panigutti, Cecilia, et al. "The role of explainable AI in the context of the AI Act." Proc. 2023 ACM Conf. on Fairness, Accountability, and Transparency. Published June 2023. See <https://dl.acm.org/doi/pdf/10.1145/3593013.3594069>
- [04] EUP approved compromise AI Act of 11.05.2023. See https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf
- [05] "AI Actors Across the AI Lifecycle Revised NIST AI RMF Figure 2". See <https://www.nist.gov/system/files/documents/2022/11/16/The%20Center%20for%20Inclusive%20Change%20-%20Attachment%20B.pdf>
- [06] OECD (2022), "OECD Framework for the Classification of AI systems", OECD Digital Economy Papers, No. 323, OECD Publishing, Paris, <https://doi.org/10.1787/cb6d9eca-en>
- [07] EUR-Lex Glossary, Trilogue. See <https://eur-lex.europa.eu/EN/legal-content/glossary/trilogue.html>
- [08] Golpayegani Delaram, Harshvardhan J. Pandit, and Dave Lewis. 'To Be High-Risk, or Not To Be—Semantic Specifications and Implications of the AI Act's High-Risk AI Applications and Harmonised Standards'. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, 905–15. FAccT '23. New York, NY, USA: Association for Computing Machinery, 2023. <https://doi.org/10.1145/3593013.3594050>
- [09] DIRECTIVE (EU) 2022/2557 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 December 2022 on the resilience of critical entities and repealing Council Directive 2008/114/EC. See <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022L2557>
- [10] Analysis of the preliminary AI standardisation work plan in support of the AI Act. JRC Report. 17.05.2023. See <https://publications.jrc.ec.europa.eu/repository/handle/JRC132833>
- [11] Treaty on the Functioning of the European Union (TFEU); See <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:12012E/TXT:en:PDF>
- [12] NIST AI 100-1 - Artificial Intelligence Risk Management Framework (AI RMF 1.0). Published Jan. 2023. See <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>
- [13] Documenting High-risk AI: A European Regulatory Perspective. IEEE Computer, Vol.56(5), May 2023. See <https://ieeexplore.ieee.org/document/10109295>
- [14] Rishi Bommasani, Kevin Klyman, Daniel Zhang, Percy Liang. "Do Foundation Model Providers Comply with the EU AI Act?". Published June 2023. Accessed at <https://crfm.stanford.edu/2023/06/15/eu-ai-act.html>
- [15] "NANDO (New Approach Notified and Designated Organisations) Information System". See <https://webgate.ec.europa.eu/single-market-compliance-space/#/notified-bodies>
- [16] "Single market and standards, Single market for goods, Building blocks of the single market, Notified bodies" See https://single-market-economy.ec.europa.eu/single-market/goods/building-blocks/notified-bodies_en

During the creation of this introductory document,
the following persons contributed, or were consulted.

AGULUCAR, ALEXEI GRINBAUM, ANDRE MEYER-VITALI, ARTHIT SURİYAWONGKUL, CHOKRI
MRAIDHA, DANIEL ALONSO, EDOARDO CELESTE, EMMANUEL KAHEMBWE, FATEMEH AHMADI
ZELETI, FRANCESCA PRATESI, J AHERN, LINDSAY FROST, MEERI, FERNANDO MORENO, NIKOLAOS
MATRAGKAS, PAOLETTO BARATTINI, RANGANAI CHAPARADZA, RAY WALSHE, SHARON FARRELL,
SILVANA MACMAHON, SONJA ZILLNER, Z AJANOVIC



adra-e.eu



[@Adra_eEU](https://twitter.com/Adra_eEU)



[Adra-e](https://www.linkedin.com/company/adra-e)

Thank
you!

