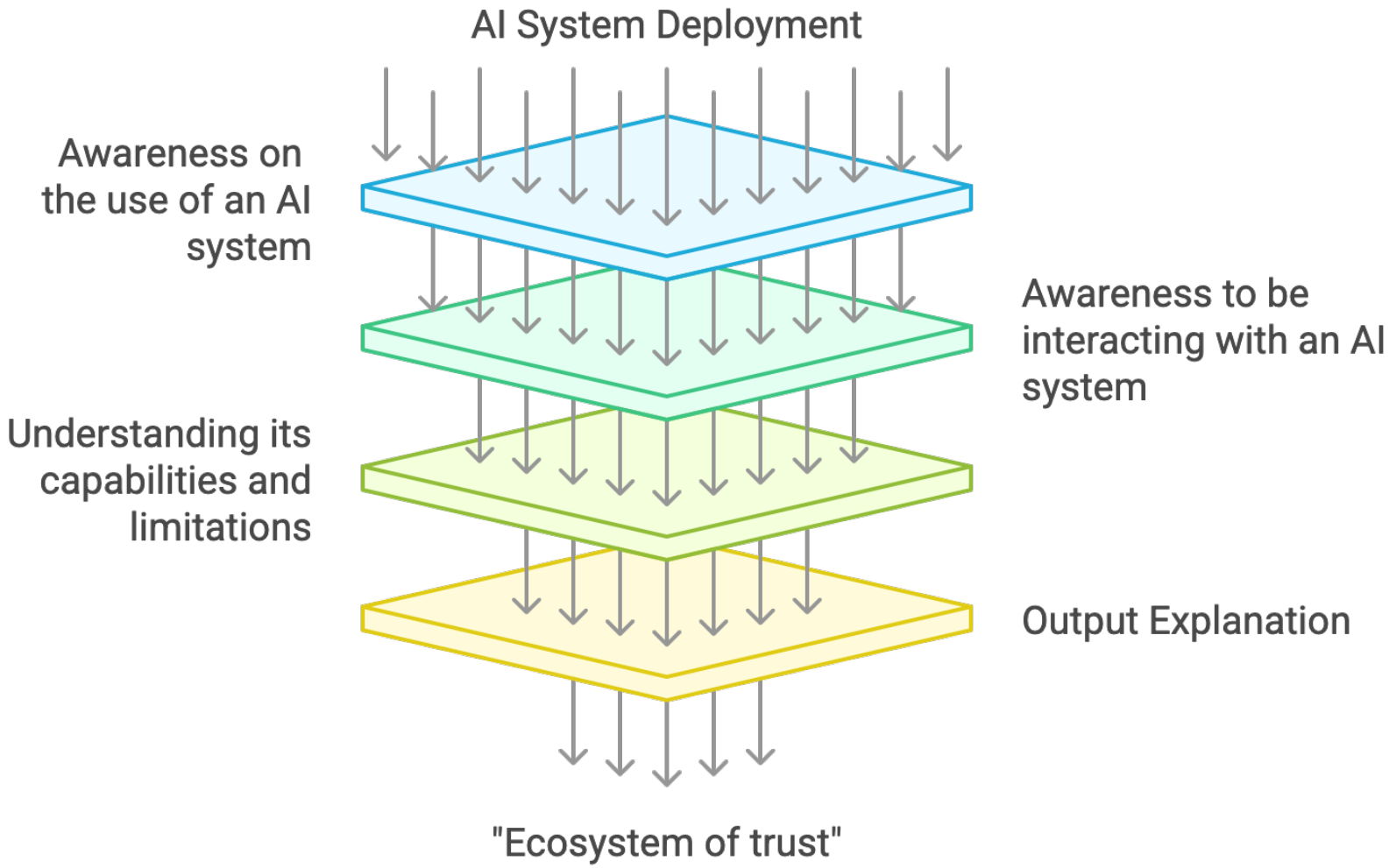
Layers of AI transparencies: opportunities and challenges in the AI Act

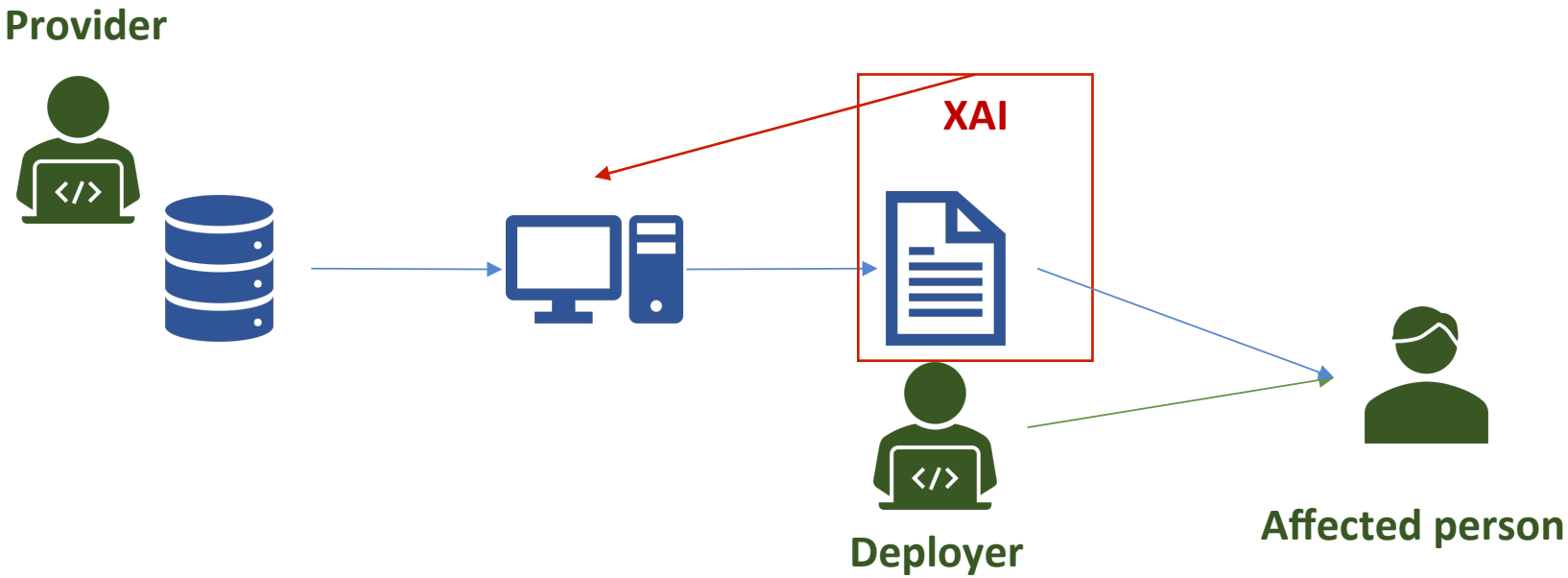
Giulia Olivato
University of Trento, FBK
golivato@fbk.eu

PON RICERCA E INNOVAZIONE 2014 - 2020
UNIVERSITÀ DI TRENTO Facoltà di Giurisprudenza
BioLaw Laboratory

Layers of AI transparencies - intro



The European Parliament's proposal



Scope

The roles of transparency(ies) in the AI Act

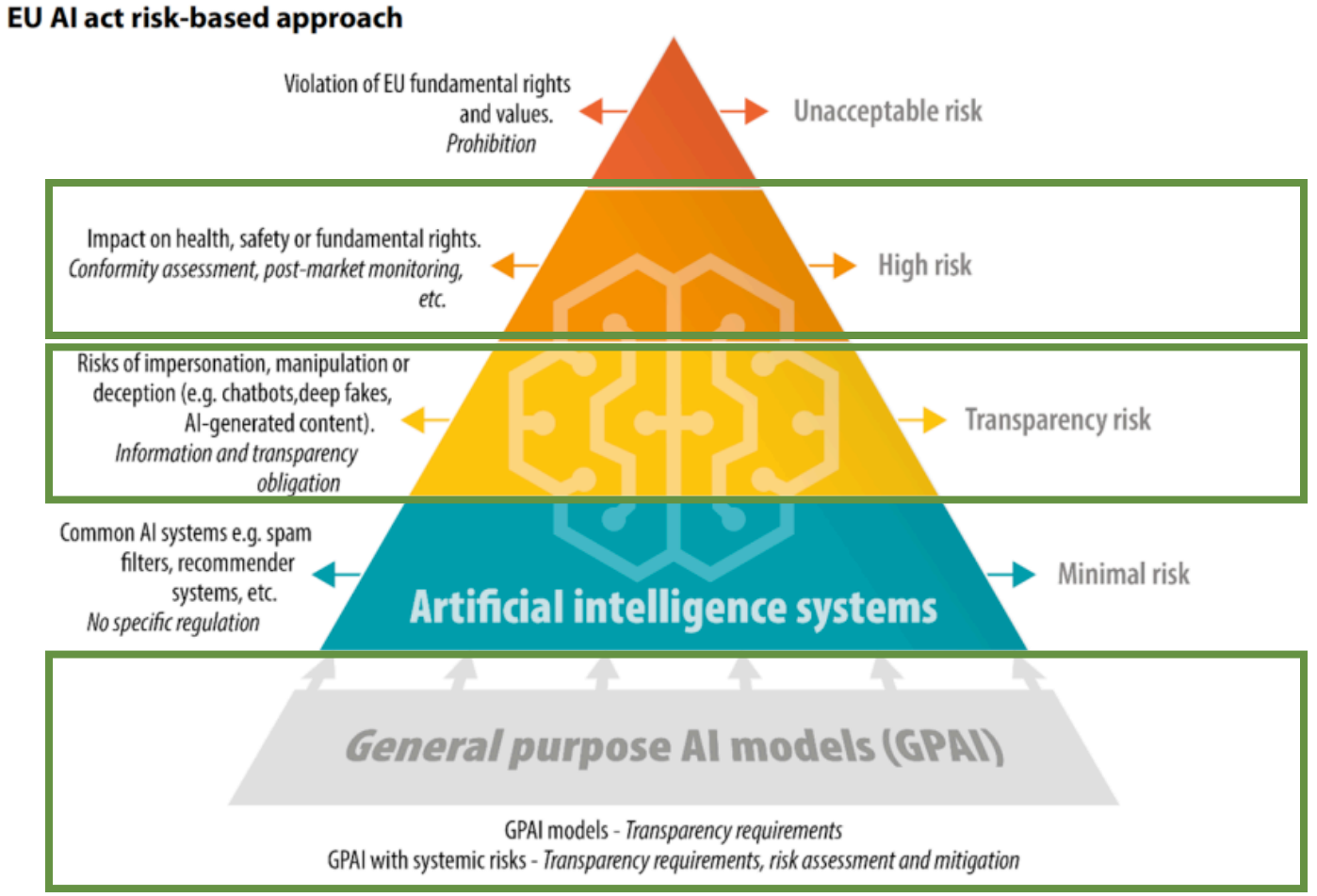
- Which transparency?
- For what purpose?
- For whom?

Transparency in the AI Act

Transparency-related provisions are present in many of the AI Act's risk levels

The AI Act has a multi-pronged and cumulative approach.
Provisions are scattered throughout the text and

- present a different scope,
- require different obligations,
- apply to different operators and
- are directed towards different AI actors

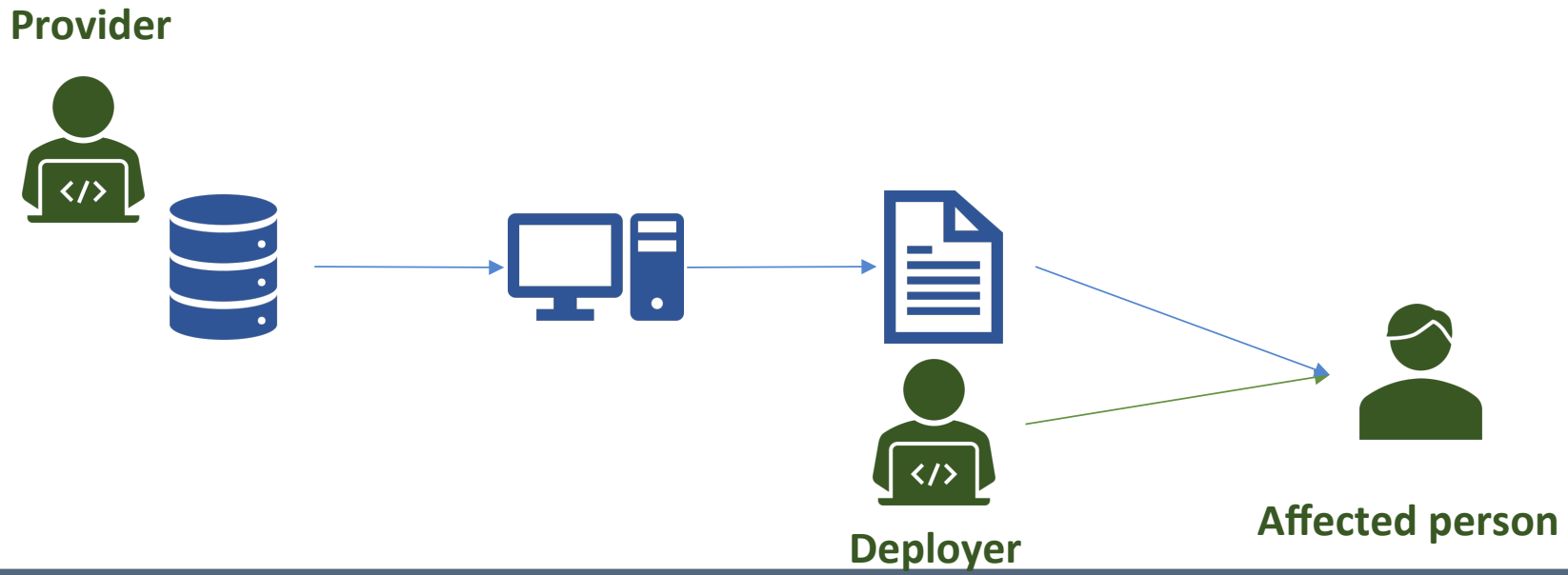


Data source: European Commission

The European Parliament's proposal

«'transparency' means that AI systems shall be developed and used in a way that allows appropriate traceability and explainability, while making humans aware that they communicate or interact with an AI system as well as duly informing users of the capabilities and limitations of that AI system and affected persons about their rights».

European Parliament, Art. 4a



What transparency?

- (A) Presence and use of an AI system
- (B) Nature of the AI systems interfacing with an individual and nature of the output
- (C) *Ex ante* information duties
- (D) Right to an explanation (?)
- (E) Information on data used by general-purpose AI models

(A) Presence and use of an AI system

Why?
Public trust

Registering high-risk AI systems in the AI Database;

- limited information in the areas of law enforcement, migration, asylum and border control management
- Only Annex III and art 6(3) AI systems

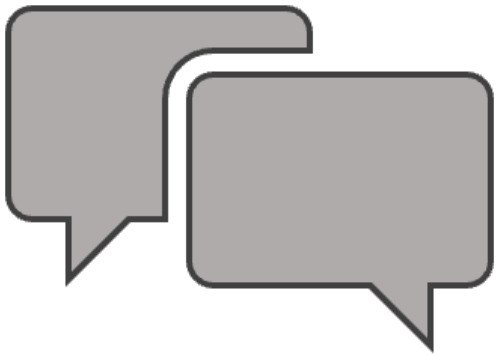
Provide information to workers on the use of AI systems in the workplace;

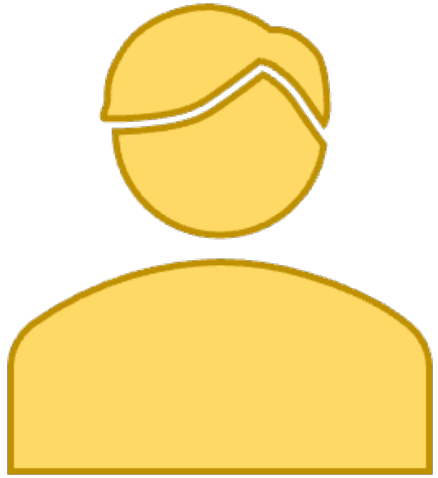
Provide information on Annex III high-risk systems in decision-making processes regarding affected individuals;

(B) Nature of the AI interlocutor (and output)

Why?
Individual trust

Article 50
Transparency obligations for providers and deployers of certain AI systems





(C) *Ex ante* information duties

Art. 13: *Trasparency*
Opening the black box? No

Transparency is instrumental to deployers being able to «interpret the system's output and use it appropriately»

Why?
Functionalist view

Target = deployer, no end users, (eg GDPR artt 14-15)

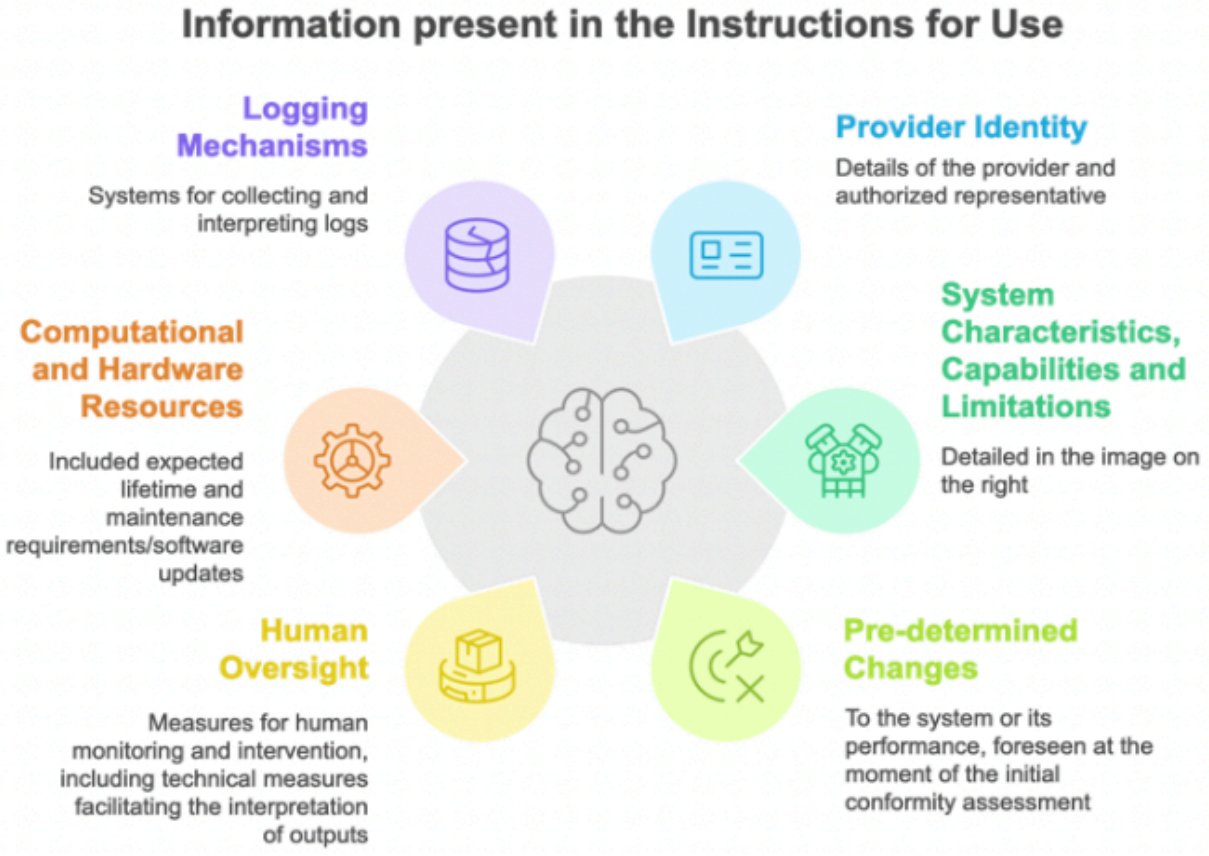
AI systems have to be designed and developed so that «their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately» and can be «effectively overseen by natural persons during the period in which the AI system is in use»

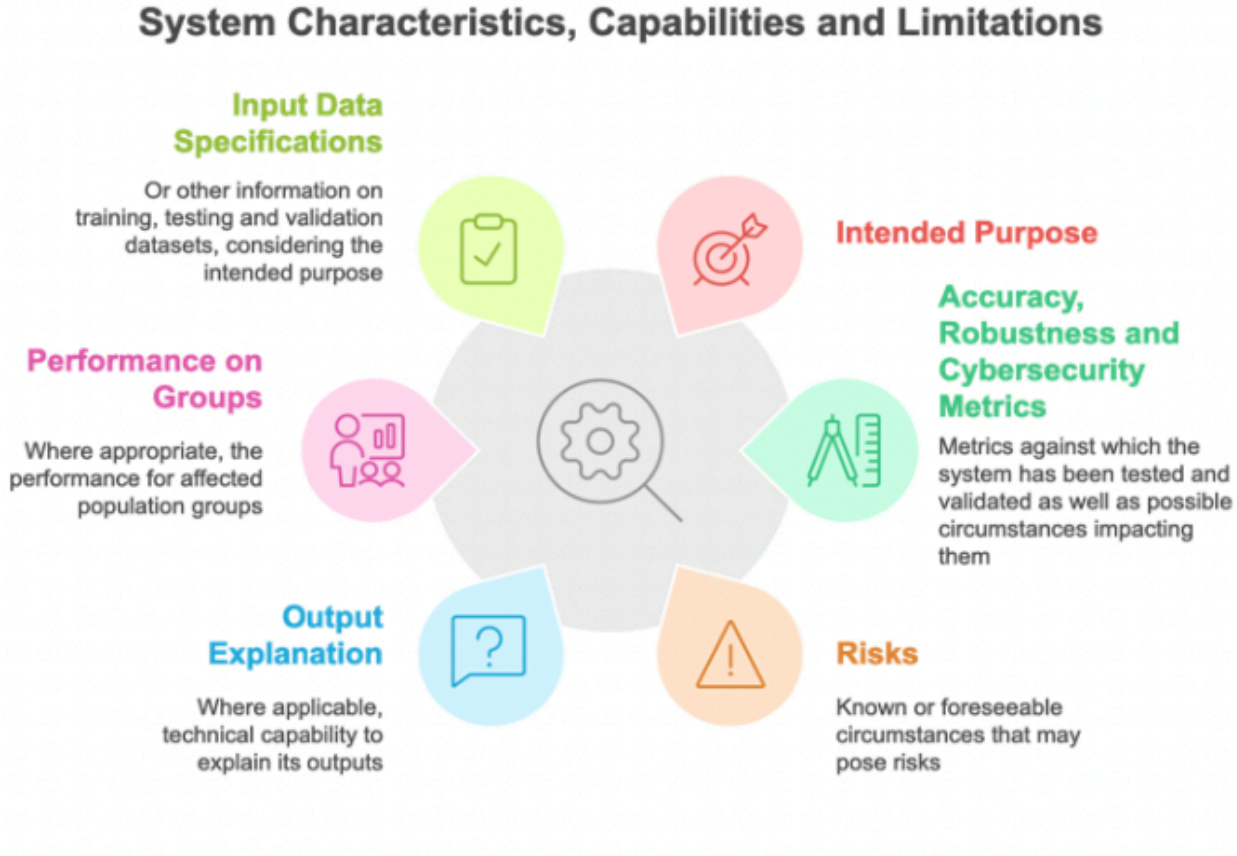
AI systems should come with 'instructions for use' which are «an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to deployers»

→ Heavily mediated information

(C) *Ex ante* information duties

Instructions for use (Art. 13)





(D) Right to an explanation?

Why?
Accountability (and contestability?)

Data subjects should always be informed when their data is used for AI training and / or prediction, of the legal basis for such processing, general explanation of the logic (procedure) and scope of the AI-system. In that regard, individuals' right of restriction of processing (Article 18 GDPR and Article 20 EUDPR) as well as deletion / erasure of data (Article 16 GDPR and Article 19 EUDPR) should always be guaranteed in those cases. Furthermore, the controller should have explicit obligation to inform data subject of the applicable periods for objection, restriction, deletion of data etc. The AI system must be able to meet all data protection requirements through adequate technical and organizational measures. ***A right to explanation should provide for additional transparency.***

EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), §60.

(D) Right to an explanation?

Why?
Accountability (and contestability?)

Article 86 §1- Right to explanation of individual decision-making

Scope

Any **affected person subject to a decision which is taken by the deployer** on the basis of the **output from a high-risk AI system listed in Annex III**, with the exception of systems listed under point 2 thereof, and which **produces legal effects or similarly significantly affects that person in a way that they consider to have an adverse impact on their health, safety or fundamental rights** shall have the right to obtain from

Explanation

the deployer **(A) clear and meaningful explanations of (B) the role of the AI system in the decision-making procedure and (C) the main elements of the decision taken.**

Article 22 §1 - Automated individual decision-making, including profiling

The **data subject** shall have the right not to be subject to a **decision based solely on automated processing**, including profiling, which produces **legal effects concerning him or her or similarly significantly affects him or her.**

Articles 86 AI Act and 22 GDPR

Art. 86's scope is concurrently narrower but broader.

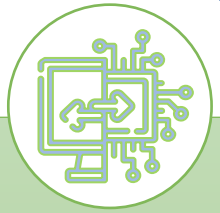
Narrower scope	
Adverse impact: only negative outcomes	All automated decision-making
Decisions affecting "health, safety or fundamental rights	All automated decision-making

Broader scope		
Scope	Hydraulic mechanism: it applies also to 'semi-automatic decisions', i.e. where a deployer uses AI output as a decisive but non-exclusive factor in decisions. It is a better reflection of AI's role in sociotechnic context; still, there can still be grey areas	Solely automated decision-making, included situations where the human rubberstamps the output (see GDPR Guidelines byEDPB and Schufa Case)
Target	Affected persons – any individual, therefore also about decisions on groups: any impacted individual	Data subjects

(E) Requirements for GPAI models

Why?
Highlighiting data use practices

Redazione e aggiornamento della **documentazione tecnica** (All. XI), che verrà trasmessa alle autorità



Attuazione di politiche e procedure (anche automatizzate) per adempiere alla normativa dell'Unione in materia di **diritto d'autore**

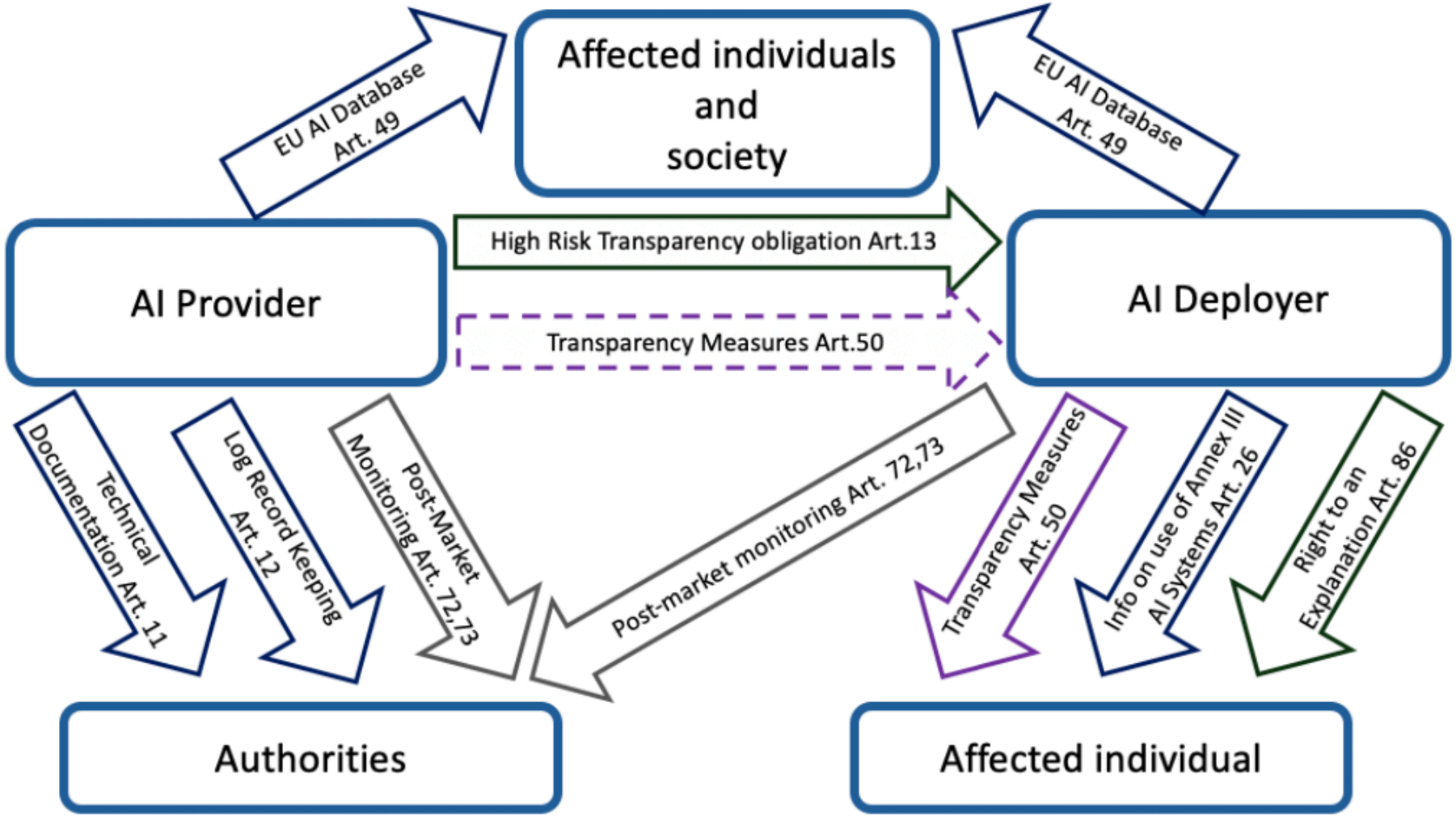
Redazione e aggiornamento di **documentazione tecnica** (All. XII) **per i fornitori a valle** di sistemi di IA che intendono integrare il modello

Redazione e messa a disposizione del pubblico di una **sintesi** sufficientemente dettagliata dei **contenuti utilizzati per l'addestramento**

Mapping transparencies

	Article	Operator subject to the obligation	Recipient of the information
Technical documentation and log record-keeping - High risk AI system	Article 11	Provider	National authorities
Transparency and instructions for use - High risk AI system	Article 13	Provider	Deployer
Use of AI systems - Annex III high risk AI systems used in the decision-making process	Article 26.7	Deployer	Affected individuals
Use of AI systems in the workplace	Article 26.11	Deployer	Workers
Presence and use of an AI system - High risk AI system	Article 49	Provider and deployer	General public
Post-market monitoring – high risk AI systems	Articles 26, 72, and 73	Provider and deployer	National authorities
Nature of the output and of the interlocutor – limited risk	Article 50	Provider and deployer	End-user
Transparency information on the data utilised	Article 53	Providers of general-purpose AI models	Downstream operators, National authorities, general public

Interrelations between provisions and actors



Conclusions

Issues:	Uncoordinated and sometimes inaccurate use of the term ‘transparency’
	Provisions are scattered and complex to operationalize together
	The AI AI Act poses particular attention to the different actors in the AI value chain (including, broadly, society and authorities) and the flow of information among them
	The AI Act requires transparency but does not specify what level is suitable for different applications, tasks or decision-making process.
	Transparency is not a goal in itself. Alone, is not enough to safeguard fundamental rights

Thank you for your attention

Giulia Olivato
University of Trento, FBK
golivato@fbk.eu