

Transparency

The fourth of the seven requirements on [Ethics Guidelines for Trustworthy AI](#) refers to “Transparency”.

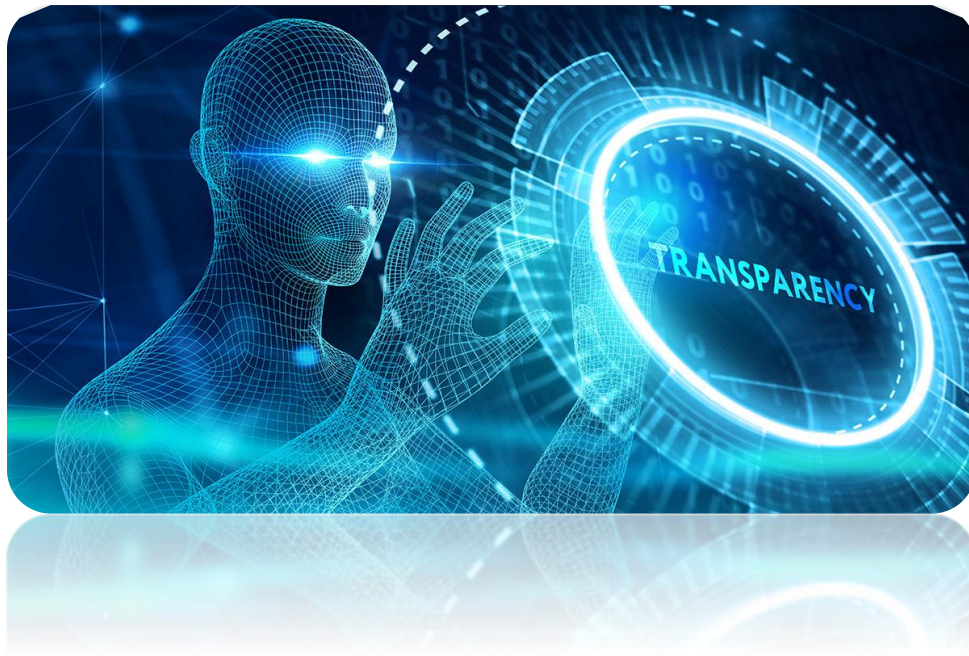


Figure 1: Transparency in an AI System

In brief, the data, AI system and AI business models should be **transparent**. **Traceability** mechanisms can help achieving this. Moreover, AI systems and their decisions should be **explained** in a manner adapted to the stakeholder concerned. **Humans need to be aware that they are interacting with an AI system and must be informed of the system’s capabilities and limitations.**

Assessing the requirement of Transparency in AI systems involves a multi-faceted approach to ensure that **AI systems operate in a clear, understandable, and accountable manner**. Transparency is crucial for fostering trust and ensuring that AI systems are used **responsibly**. Critical aspects for addressing transparency include the following:

Documentation

AI systems should have **comprehensive documentation detailing their design**, purpose, data sources, decision-making processes, outcomes, and potential impacts. This documentation should be accessible to various stakeholders, including developers, (end-)users, and regulators.

Explainability

The AI system should provide simple, clear explanations for its decisions and how outcomes were produced. Explanations should be provided in **a manner that is**

understandable to non-experts so that users can understand, and therefore trust, the decision made by the AI system.

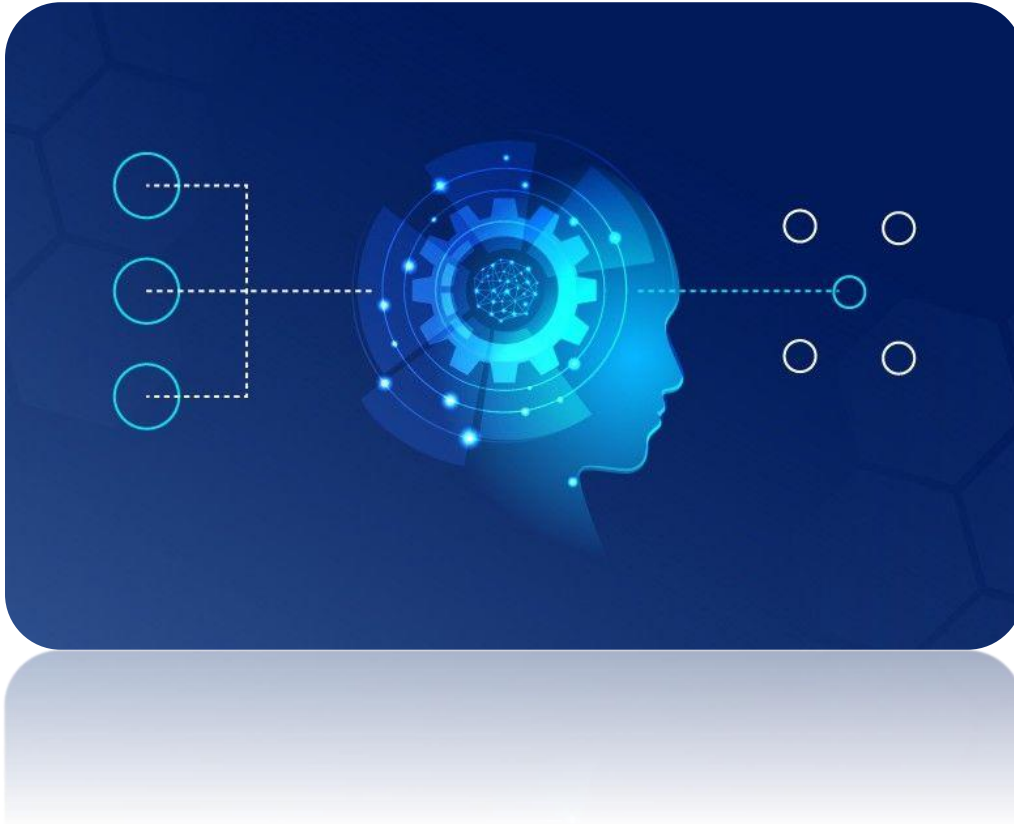


Figure 2: Explainability in an AI System

Users should also have access to information on the degree to which the AI system influences human and organisational decision-making processes, including the rationale, data inputs, and key decision criteria so that they are enabled to go beyond understanding into taking appropriate actions if needed.

Disclosure

The existence of AI systems and their purposes must be **disclosed to users**, including sufficient communication regarding the interaction with an AI system and not with other humans. In addition, provision of additional simple and clear information should be ensured, covering how decisions are made by AI, what data is used, the underlying logic or algorithms employed, but also limitations and potential shortcomings.

Transparency in MANOLO

Within MANOLO, technical partners will closely work with **FOUR DOT INFINITY INFORMATION AND TELECOMMUNICATIONS SOLUTIONS**, to deliver explainable AI techniques and mechanisms that will ensure that all AI systems' decisions and outcomes are understandable by end-users without requiring a detailed knowledge of the underlying processes.

Dedicated **socio-technical scenarios will be drafted following the Z-inspection® process** (under the guidance of the [Arcada University of Applied Sciences](#)), to ensure that the expected degree of explainability is achieved.

WRAP UP

To achieve transparency in AI systems a thorough evaluation of documentation, explainability, and disclosure of information should be ensured, complementing the other key requirements for Trustworthy AI.

Users should be aware that they are interacting with an AI system, including information about its rationale, data, capabilities and limitations. AI decisions and outcomes must be explained in a way that is easy for users to understand. By addressing these areas, we can ensure that AI systems are transparent, fostering trust and accountability among users and stakeholders.

In the MANOLO, partners are collaborating to ensure transparency by implementing explainable AI techniques, making AI decisions understandable to users. Clear socio-technical scenarios will also support the implementation of AI systems that meet transparency requirements, building trust and ensuring responsible use of AI, compliant with relevant standards, regulations and guidelines.

Learn more about MANOLO EU Project here: <https://manolo-project.eu/>