

AI Trustworthiness in the Cloud Edge Continuum – The MANOLO Approach

In December 2023, the European Parliament and the Council of the EU reached a political agreement on the AI Act. This landmark legal framework introduces a risk-based approach (Figure 1) to regulating AI systems based on their potential impact on human rights and society.

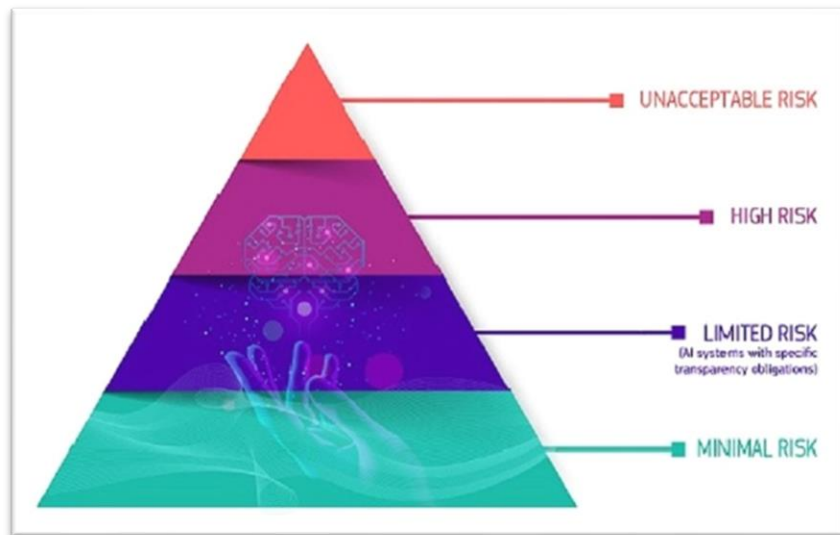


Figure 1: The 4 levels of risk for AI systems within the AI Act [source]

As the AI Act aims to ensure that Europeans can trust what AI has to offer, the framework builds on the concept of trustworthiness to safeguard that AI systems operate as intended and do no harm physically or ethically to their users.

Trustworthy AI in Europe

Based on the Ethics Guidelines for Trustworthy AI defined by the EU High-Level Expert Group on AI, Trustworthy AI (TAI) systems should be lawful, ethical, and robust throughout their entire life cycle.

The foundations of TAI lay upon 4 ethical principles, to ensure ethical and robust AI:

1. Respect for human autonomy
2. Prevention of harm
3. Fairness

4. Explicability

Furthermore, to realise TAI systems, the aforementioned ethical principles are translated into 7 (seven) interconnected key requirements, of equal importance (Figure 2), that AI systems should implement and meet throughout their entire life cycle:

1. Human Agency and Oversight
2. Technical Robustness and Safety
3. Privacy and Data Governance
4. Transparency
5. Diversity, non-discrimination, and fairness
6. Societal and Environmental well-being
7. Accountability



Figure 2: Interrelationship of the 7 Seven Requirements.

MANOLO adopts a trustworthiness-by-design approach by employing the Z-Inspection® process to ensure that all AI solutions implemented will be trustworthy from design to

deployment. On top of the 7 (seven) requirements introduced by the EU, the Z-Inspection® Initiative brings forth two additional requirements:

8. Respect the values of Western European Democracy
9. Use of the AI results in concentration of power

The Z-Inspection® process

The Arcada University of Applied Sciences, a partner in the MANOLO consortium and affiliated lab to the Z-Inspection® Initiative, is supporting the Trustworthy AI co-design using the Z-Inspection® process (Figure 3), to ensure that MANOLO's AI-based solutions and their application in the envisioned use cases are aligned with the trustworthy requirements, delivering clear evidence along their implementation.

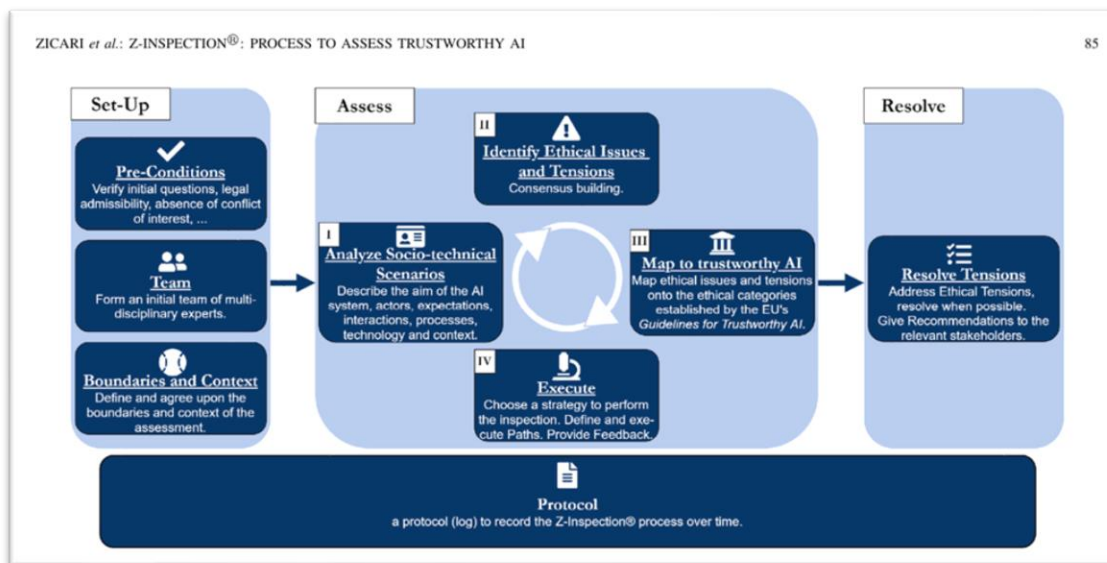


Figure 3: Z-inspection process in a nutshell.

Z-Inspection® in the EU-funded MANOLO project.

Starting with training that took place on the 14th of March 2024, the MANOLO consortium became accustomed to the Z-Inspection® process and started working together with ARCADA in setting up the steps to be followed during the co-design of the MANOLO architecture and use cases' deployment plans.

With a focus on the use cases, the consortium will delve into the assessment phase, collaborating on (i) defining the socio-technical scenarios of the overall MANOLO framework, (ii) identify ethical issues and tensions, (iii) map with the abovementioned key requirements, and (iv) execute an in-depth assessment of each component, the overall

system, and their impact to the MANOLO use cases. Moreover, the socio-technical scenarios will be deployed under cloud, edge, and cloud-edge continuum use cases. In this way, the consortium aims to provide comprehensive insights into the practical implications and benefits of MANOLO across diverse deployment environments, ensuring its adaptability and effectiveness in real-world settings.

Key insights from the application of Z-Inspection® in the project will be published in several scientific venues (conferences, journals, etc.). The main takeaways will feed the delivery of the MANOLO Code of Conduct that will aim to act as a guideline for European SMEs into developing and deploying efficient and trustworthy AI solutions, aligned with the EU values and principles, and of course the AI Act.

Wrap Up

To sum it up, MANOLO aims to deliver a set of trustworthy-AI-by-design solutions for the cloud-edge continuum, adopting the Z-Inspection® process to guide, support, assess, and validate each step and final outcome. Through the application of rigorous assessment methodologies, MANOLO not only contributes to advancing the field of trustworthy AI but also provides valuable guidance in navigating the complexities of AI development and deployment in accordance with regulatory frameworks and ethical considerations.

Stay tuned, to learn more about this journey and explore along the MANOLO consortium the development of trustworthy-AI-by-design.