8

Conclusion

This work explores functional and non-functional requirements for edge AI systems, providing a comprehensive framework for understanding and implementing these advanced technologies across the micro-, deep-, and meta-edge continuum. It provides a holistic approach to edge AI system development and deployment by intertwining the concepts of dependability and trustworthiness with these requirements.

The shift towards edge AI represents a significant evolution in AI technology, offering solutions to challenges posed by centralised systems, such as latency, bandwidth limitations, and privacy concerns. However, this transition also introduces new complexities and considerations that must be carefully addressed to ensure the successful implementation of edge AI systems.

The detailed examination of functional requirements highlights the critical capabilities for edge AI systems to perform effectively in diverse, often resource-constrained environments. These include real-time processing, energy efficiency, and adaptive learning capabilities. Equally important are the non-functional requirements, which encompass a wide range of quality attributes from reliability and security to more AI-specific considerations like explainability and fairness.

By linking these requirements to the concepts of dependability and trustworthiness, we underscore the importance of creating edge AI systems that are functionally effective, reliable, secure, and aligned with ethical standards and societal values. This approach recognises that the success of edge AI deployments hinges not just on technical performance but also on the ability to build and maintain trust with users and stakeholders.

Discussing KPIs and measurement methods provides practical guidance for assessing and monitoring edge AI systems. This focus on quantifiable metrics and ongoing assessment is crucial for maintaining system performance and adapting to changing requirements over time.

The framework presented in this work offers a foundation for future research and development in edge AI. It highlights areas where further innovation is needed, particularly in addressing the unique challenges of edge environments, such as limited computational resources and intermittent connectivity.

Moreover, the emphasis on ethical considerations and alignment with societal values points to the need for ongoing dialogue between technologists, policymakers, and the public. Ensuring their trustworthiness and societal acceptance will be important as edge AI systems become more prevalent in critical applications.

The international standardisation landscape for AI is dynamic and rapidly maturing. Led by key organisations like ISO/IEC, IEEE, ITU-T, ETSI, and CEN-CENELEC, a comprehensive framework of standards is emerging to guide the responsible development and deployment of AI and edge AI technologies. These standards provide essential guidance on everything from foundational concepts and data quality to the critical issues of trustworthiness, ethics, and risk management.

While significant progress has been made, the work is far from complete. The continued evolution of AI technology, particularly in areas like generative AI and edge AI, will require ongoing development and harmonisation of standards. Collaboration between the various standards bodies will be crucial to ensure a cohesive and globally relevant framework that can foster innovation while safeguarding against potential harms. For organisations developing or deploying AI, active engagement with and adoption of these standards is both best practice and a strategic imperative for building trust, ensuring compliance, and unlocking the full potential of AI and edge AI.

The comprehensive exploration of edge AI requirements provides a valuable resource for researchers, developers, and decision-makers in the field. Offering a structured approach to understanding and implementing edge AI systems contributes to the responsible advancement of this continuously changing technology. As edge AI continues to evolve, the principles and frameworks outlined in this work will serve as essential guides for creating systems that are not only technologically advanced but also dependable, trustworthy, and aligned with broader societal goals.

This publication is a detailed written study focused on functional and non-functional requirements for edge AI systems to provide an overview of current state of play, concepts, definitions, taxonomy and gaps in the field.

Acknowledgements

This publication has received funding through the projects Horizon Europe dAIEDGE [1] and Chips JU EdgeAI [3]. The Horizon Europe dAIEDGE "A network of excellence for distributed, trustworthy, efficient and scalable AI at the Edge" project is supported under grant agreement No 101120726. The Chips JU EdgeAI "Edge AI Technologies for Optimised Performance Embedded Processing" project is supported by the Chips Joint Undertaking and its members including top-up funding by Austria, Belgium, France, Greece, Italy, Latvia, Netherlands, and Norway under grant agreement No 101097300. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Chips Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them.