

Trustworthy Impact Assessment of a Federated Machine Learning System in Healthcare Product Development

Antti Happonen and Mark van Gils
Decision Support for Health

Health Informatics, Public Health, Trustworthy AI, AI Impact Assessment, Z-Inspection

Abstract

Federated machine learning (FML) is based on machine learning or deep learning methods, where models are trained locally using distributed data sources and only the parameters of the machine learning results are aggregated instead of the raw data itself [1]. The FML method is particularly well suited for healthcare, where patient privacy, data sensitivity, and regulatory compliance are paramount.

The *High-Level Expert Group on Artificial Intelligence* (AI), established by the *European Commission*, defines ethical guidelines for trustworthy AI [2]. These ethical guidelines describe seven key principles that developers must follow to ensure the ethical and responsible development of AI. Trustworthy AI refers to systems that are explainable, fair, interpretable, robust, transparent, safe and secure.

We introduce the *Federated Ethical Healthcare Learning Sandbox* project funded by *Nordic Innovation*. Our project demonstrates how FML can be trustworthily applied to unstructured clinical data across the Nordic region.

Z-Inspection® provides a methodology for assessing AI trustworthiness in real-world settings [3]. We implemented independent co-design processes involving developers, physicians, and Z-Inspection® assessors for two development teams. The developed AI systems were integrated into an FML architecture aimed at generating personalized treatment predictions for migraine or headache patients. Through Z-Inspection®, we have demonstrated that the FML architecture supports both data sovereignty and ethical consistency across jurisdictions.

Our results provide insight into how EU legislation on ethical AI and healthcare product development can be integrated into the regulatory context and how Z-Inspection® assessment could be an integral part of the management processes of AI-based product development projects.

References

- [1] Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*.
- [2] Ala-Pietilä, P., Bonnet, Y., Bergmann, U., Bielikova, M., Bonefeld-Dahl, C., Bauer, W., ... & Van Wynsberghe, A. (2020). The assessment list for trustworthy artificial intelligence (ALTAI). *European Commission*.
- [3] Zicari, R. V., Brodersen, J., Brusseau, J., Düdder, B., Eichhorn, T., Ivanov, T., ... & Westerlund, M. (2021). Z-Inspection®: a process to assess trustworthy AI. *IEEE Transactions on Technology and Society*, 2(2), 83-97.