Aligning XAI with EU Regulations for Smart Biomedical Devices: A Methodology for Compliance Analysis

Francesco Sovrano[1][0000-0002-6285-1041], Michael Lognoul[2][0009-0005-5137-8278] and Giulia Vilone [0000-0002-4401-5664]

[1] Università della Svizzera italiana, Via la Santa 1, 6900 Lugano, Italy
[2] University of Namur (CRIDS-NADI), Rue de Bruxelles 61, 5000 Namur, Belgium

Keywords: XAI, Legal XAI, Compliance with explanation requirements, medical devices.

Abstract

Context, issue and objective. Incorporating AI into medical devices represents a tremendous leap in medical technology: this integration results in substantial improvements in patient care [1]. A key evolution is the shift from open-loop systems, where physicians interpret data to inform decisions, to more sophisticated closed-loop and semi-closed-loop systems, where devices autonomously or semi-autonomously adjust their operations based on continuous monitoring. However, a critical challenge with such advanced AI systems in healthcare is their 'black box' nature [2].

In parallel, several EU regulatory frameworks require transparency and the provision of explanations in various situations, including when medical devices are at stake, to ensure AI technologies' transparency, fairness, and accountability [3]. For instance, these devices must comply with the General Data Protection Regulation (GDPR), the Artificial Intelligence Act (AIA), and the Medical Devices Regulation (MDR), each contributing unique requirements related to transparency and explainability. This is where Explainable Artificial Intelligence (XAI) plays a crucial role, offering tools to help comply with the legal framework and make the inner workings of complex systems more understandable to the diverse stakeholders involved in their operation [4].

However, navigating the complex regulatory landscape poses significant challenges for developers and researchers. Selecting and implementing the correct XAI algorithms to reach compliance with EU regulations is a major hurdle [5] [6]. To help tackle this issue, we carried out a thorough analysis of various XAI algorithms to determine if they can help satisfy explainability requirements set by the GDPR, the AIA, and the MDR. To this end, we have developed a novel methodology, to evaluate and understand the role of XAI in adhering to the legal frameworks studied.

Main contribution: a novel methodology. Our methodology, summarized by Fig. 1, comprises the following five steps. **1.** Adoption of Explanation and Explainability Definitions. To map XAI methods with legal explanatory requirements, we adopted the definition formalized by Sovrano and Vitali [7], which conceptualizes explanations as answers to questions that produce understanding. According to this definition, an

explanation provides sufficient information for an audience's understanding. **2.** Legal Analysis of Explanation Requirements. A legal expert thoroughly analyzed the GDPR, AI Act, and MDR to pinpoint their explanation requirements and characteristics. Then, following an inductive coding approach [8], the legal expert identified the high-level explanatory goals underlying these requirements. **3.** Identification and Classification of XAI Methods. Concurrently, two AI experts conducted a literature review to compile a comprehensive (but not exhaustive) list of existing XAI methods and categorized them based on their explanation format, input format, and model-agnostic status to discern the types of explanatory questions they could address. We then used a question-driven design process in which XAI methods are mapped to explanatory questions based on their characteristics [9]. **4.** Aligning XAI Methods with Legal Requirements. By performing a deductive thematic analysis [8], we mapped the XAI questions to the legal explanatory goals enshrined in the GDPR, AIA, and MDR. This was possible precisely because the adopted working definition of explanation is 'answers to questions'. Eventually, we could identify congruence where XAI capabilities can be exploited to help meet the stipulated legal explanation requirements. This matching process ensures that the selection of XAI methods is technologically sound and legally robust. **5.** Case Studies. We ultimately illustrated our findings with two use cases in the field of medical devices incorporating complex, black-box AI systems.

Beyond the proposed matching between (i) explanation requirements arising from GDPR, AI ACT or MDR and (ii) XAI methods, we developed a set of instructions to aid developers and researchers in selecting the most appropriate XAI algorithms for different medical devices. These instructions and the methodology developed in the paper provide a fundamental framework designed to be flexible and seamlessly incorporate newly emerging XAI algorithms and evolving regulations.

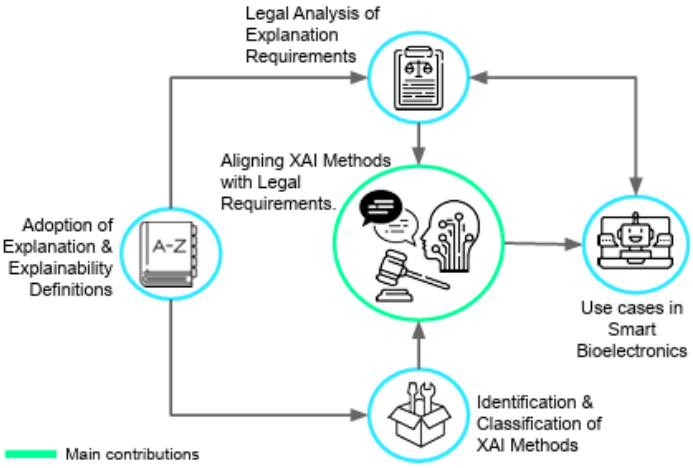

Fig. 1. Schematic overview of our research methodology for integrating legal requirements and XAI tools.

Full paper available at: https://ebooks.iospress.nl/doi/10.3233/FAIA392 (p. 826-833)

This paper was originally submitted to ECAI2024 and published in the proceedings of the conference.

Acknowledgments. F. Sovrano acknowledges the support of the Swiss National Science Foundation for the SNF Project 200021_197227. M. Lognoul acknowledges the support of the European Union H2020 research and innovation program, Grant Agreement No. 958339 (DENiM), and the support of the ARIAC by Digital Wallonia4.AI (SPW Recherche/Wallonia Research) project, Grant agreement No. 2010235.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this abstract, nor to the content of the paper discussed.

References

1. C. Guger, N. F. Ince, M. Korostenskaja, and B. Z. Allison. Brain computer interface research: A state-of-the-art summary 11. In Brain Computer Interface Research: A State-of-the-Art Summary 11. 2024.
2. J. Bernal and C. Mazo. Transparency of artificial intelligence in healthcare: insights from professionals in computing and healthcare worldwide. Applied Sciences, 2022.
3. A. Bibal, M. Lognoul, A. De Streel, and B. Frénay. Legal requirements on explainability in machine learning. Artificial Intelligence and Law, 2021.
4. F. Sovrano and F. Vitali. Explanatory artificial intelligence (yai): human-centered explanations of explainable ai and complex data. Data Mining and Knowledge Discovery, 2022.
5. K. M. Richmond, S. M. Muddamsetty, T. Gammeltoft-Hansen, H. P. Olsen, and T. B. Moeslund. Explainable ai and law: An evidential survey. Digital Society, 2024.
6. T. Kirat, O. Tambou, V. Do, and A. Tsoukiàs. Fairness and explainability in automatic decision-making systems. a challenge for computer science and law, 2022.
7. F. Sovrano and F. Vitali. An objective metric for explainable ai: how and why to estimate the degree of explainability. Knowledge-Based Systems, 2023.
8. J. Fereday and E. Muir-Cochrane. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International journal of qualitative methods, 2006.
9. Q. V. Liao, M. Pribi´ c, et al. Question-driven design process for explainable ai user experiences. arXiv preprint arXiv:2104.03483, 2021.