Trustworthy AI in Cancer Imaging Research

Ioanna Chouvarda • Sara Colantonio
Gianna Tsakou • Guang Yang
Editors

Trustworthy AI in Cancer Imaging Research

Springer

Editors
Ioanna Chouvarda
Laboratory of Computing
Medical Informatics and Biomedical
Imaging Technologies
School of Medicine
Aristotle University of Thessaloniki
Thessaloniki, Greece

Gianna Tsakou
Research & Development Department
MAGGIOLI SpA Greek Branch
Marousi, Greece

Sara Colantonio
National Research Council of Italy
Institute of Information Science and Tec
Pisa, Italy

Guang Yang
Bioengineering Department and Imperial-X
Imperial College London
London, UK

ISBN 978-3-031-89962-1 ISBN 978-3-031-89963-8 (eBook)
https://doi.org/10.1007/978-3-031-89963-8

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book was inspired and supported by the collaborative work in the AI for Health Imaging cluster (AI4HI). AI4HI was created as a networking initiative among, initially, five and now seven EU research and development projects focusing on the creation of big cancer image data repositories, as well as development and validation of AI models in support of decision-making of healthcare professionals for various cancers and related clinical challenges linked to various cancer imaging modalities.

The book has Trustworthy AI in Cancer Imaging as a central theme. Due to the nature of the topics addressed in these projects and discussed in the AI4HI cluster, Trustworthy AI has been an emerging and horizontal theme, which triggered the idea of creating this book in order to discuss various aspects linked to and supporting trustworthy AI. The chapter contributors, in their vast majority, are researchers from the EU-funded projects CHAIMELEON, EUCANIMAGE, INCISIVE, PRIMAGE, PROCANCER-I, RADIOVAL, and EUCAIM, combining a wide variety of disciplines, thus reflecting the multi-dimensional nature of trustworthy AI.

Cancer research has been central in the biomedical community, seeking earlier and safer cancer diagnosis and prognosis, as well as better and more personalized treatment decisions. AI is increasingly receiving attention as a major component leveraging cancer imaging and multi-omics data analysis, toward these goals. A key factor to ensure the success and impact of AI in cancer research and secure the wider adoption in clinical practice is the aspect of trustworthiness. This entails a multifaceted strategy in all phases of an AI service development, spanning from AI design to training and validation and including—among others—human oversight, technical robustness, data governance, transparency, fairness, and auditability. AI trustworthiness is currently moving from the definition of its theoretical framework capturing all trustworthiness perspectives to guidelines and best practices for its practical implementation. It is also evolving to incorporate the particularities of the cancer research areas and the needs of the whole user spectrum.

Both Cancer Imaging and Trustworthy AI receive attention in the academic and clinical societies, and their meeting point is expected to create value in cancer diagnosis and therapy. In complex, multidisciplinary, emerging fields, achieving

consensus in key research elements is crucial, as well as translating theoretical approaches to actual practice.

In this respect, the purpose of this book is to provide an end-to-end discussion on the multi-dimensional challenges related to trustworthy AI: from general principles and design considerations, to data collection, harmonization, and management, up to AI training, validation, and clinical deployment. The discussion and approach in each chapter is blending scientific know-how in each of the relevant scientific domains with findings from the practical implementation of the various multi-disciplinary approaches applied in each of the projects.

The book is organized in four parts, with three chapters per part:

- Part I: Overall Considerations
- Part II: Preparing for Trustworthy AI: The Data and Metadata for Quality, Transparency and Traceability
- Part III: Implementing Trustworthy AI: The Algorithms and DSS
- Part IV: Validating Trustworthy AI: The Validation and User Perspective

An overview of the chapters in each of the four parts follows.

Part I focuses on general design considerations and principles of trustworthy AI.

Chapter 1 discusses societal, legal, and ethical aspects of Trustworthy AI. It explores the complex interplay between ethical, legal, social, and societal dimensions that impact trustworthy AI development and deployment. The chapter analyzes but also integrates perspectives from social sciences, philosophy, ethics, and law. The interdisciplinary approach goes beyond the mere involvement of diverse areas of expertise, emphasizing a "layered" understanding that addresses all the elements of an Ethical, Legal, Societal Implications (ELSI) approach.

Chapter 2 examines clinical considerations related to trustworthy AI in oncologic imaging. The chapter zooms into the concept of trust in medicine and explores aspects influencing trust, as viewed from the clinical point of view. Such aspects encompass AI tool design, data collection and preparation, AI training testing and validation, technology assessment, clinical validity, user experience, AI robustness, explainability and interpretability, generalizability. The chapter also provides a high-level overview of the various approaches applied in the AI4HI projects.

Chapter 3 underscores the importance of stakeholder engagement and a social innovation approach for trustworthy AI by leveraging the transferable experiences from one of the AI4HI projects, the EU-funded RadioVal project on breast cancer imaging. The chapter stresses the importance of understanding the factors that lead to mistrust of AI tools among various stakeholders and emphasizes the engagement of all relevant actors from the early stages of AI design and development as a key strategy to enhance trustworthiness. The structured process and outcomes of the multi-stakeholder iterative engagement strategy adopted in the RadioVal project are presented, providing a transferable framework for the design and implementation of trustworthy AI-driven tools in healthcare.

Part II discusses the data challenges that have to be addressed as a prerequisite and a preparation step for developing trustworthy AI models.

Chapter 4 focuses on the development of trustworthy AI synthetic datasets in oncology via deep learning-based generative algorithms. It starts with an introduction of concepts and techniques on synthetic data generation in cancer research, followed by the presentation of key successes and applications found in the literature. Methods to assess quality, reliability, and trustworthiness of synthetic data are discussed, and finally challenges and opportunities are identified for synthetic data in cancer research.

In *Chapter 5*, data and metadata standardization is discussed as a means to ensure uniformity and consistency across different diagnostic modalities, enabling data use. The standardizing approaches, the design decisions, and the respective workflows adopted in three EU cancer imaging projects are illustrated, followed by a discussion on experiences and problems for future consideration.

Chapter 6 delves into the role of data harmonization in overcoming challenges in integrating diverse, multimodal, and multicenter datasets, toward facilitating the generation of comprehensive data repositories to foster research. The different steps required to build such repositories are elaborated including data de-identification, image annotation, quality control, and harmonization. Best practices and challenges are discussed.

Part III focuses on the implementation of trustworthy AI, spanning from the underlying infrastructures, the issues needed to ensure robustness to real examples in research initiatives.

Chapter 7 details the application of AI models in cancer diagnosis and prognosis, focusing on projects using various imaging modalities. The authors describe the creation of AI pipelines, emphasizing data preprocessing (including harmonization) and the use of various machine and deep learning techniques. Specific AI models developed for breast, lung, ovarian, and prostate cancer are presented, thereby highlighting their performance in lesion detection, segmentation, classification, patient stratification, and risk prediction. Challenges encountered, such as data scarcity and variability, and the importance of explainability, uncertainty estimation, and adversarial robustness are discussed.

Chapter 8 explores AI-enabling infrastructures, focusing on cloud computing, edge computing, and federated learning. It compares traditional IT systems to cloud-based architectures, highlighting the benefits of characteristics such as scalability, elasticity, and resource pooling offered by cloud solutions for AI workloads. The chapter also discusses the challenges and opportunities presented by the increasing use of AI in sensitive data domains like healthcare, emphasizing the need for privacy-preserving techniques. Hence, it delves into edge computing and federated learning as viable approaches for training AI models on decentralized data while maintaining data privacy. With these two new paradigms, AI models can be trained in a trustworthy manner maintaining the data in the original organizations with similar performance. Finally, the chapter highlights the challenges raised by these new paradigms, particularly of federated learning, which is still a very young research field and requires research expertise to be effectively employed.

Chapter 9 investigates the challenges of creating robust clinical AI for medical imaging. It argues that successful AI requires four key aspects: generalizability,

explainability, uncertainty estimation, and adversarial resistance. The authors discuss techniques to improve these aspects and address vulnerabilities to malicious attacks, emphasizing the need for both technical solutions and strong ethical considerations for responsible AI deployment in healthcare. Some key challenges, such as population representativeness, are highlighted as urgent issues.

Part IV: Real-Life Deployment of AI Services—Practical Considerations and Clinical Validation

This part explores the real-world deployment of AI in healthcare, with a focus on practical challenges, intellectual property protection, clinical validation, and regulatory pathways. The discussion encompasses key considerations such as data harmonization, interoperability, ethical compliance, and the need for rigorous validation to ensure AI-driven solutions are safe, effective, and seamlessly integrated into clinical workflows.

Chapter 10 examines the practical implications of deploying AI services, addressing the importance of regulatory approval (e.g., CE mark, FDA), patenting strategies, and accelerating clinical translation. It highlights AI's role in medical imaging, clinical decision support, and drug development, emphasizing the need for continuous model monitoring, ethical safeguards, and public-private collaborations to foster trust and adoption.

Chapter 11 underscores the importance of high-quality, harmonized data in medical repositories, particularly in cancer imaging. It discusses best practices in data pre-validation to enhance AI model reliability, drawing insights from large-scale EU projects such as INCISIVE, ProCAncer-I, and EUCAIM. These repositories support personalized medicine, early cancer detection, and AI-driven clinical decision-making by integrating multimodal imaging, genomic, and clinical trial data.

The final chapter, *Chap. 12*, focuses on the clinical validation of AI solutions, using experiential learning from the AI4HI network and its EU-funded projects. It highlights the challenges of interoperability, standardization, generalizability, and ethical compliance. Lessons learned from these initiatives provide a framework for future AI integration, ensuring robust validation and regulatory alignment to enhance clinical outcomes.

Through these chapters, the four parts provide a comprehensive roadmap for generating cancer imaging data repositories and trustworthy AI models, as well as toward successfully transitioning AI innovations from research to clinical practice while addressing regulatory, technical, and ethical challenges.

Overall, this book distils the experience and consensus from big projects in the domain of Cancer Imaging Research, including best practices toward practical implementation of trustworthy AI, which can be useful for researchers interested in AI in Medical Imaging and Health Sciences with a multidisciplinary perspective. The book can be a reference point for Digital Skills in Radiology and/or Cancer Imaging Research, and beyond.

This has been the collective effort of numerous authors. Together with the huge effort and contribution of the chapter authors, we wish to also acknowledge all the partners who contributed to nurturing discussions and collaboration in the AI4HI cluster.

Thessaloniki, Greece — Ioanna Chouvarda
Pisa, Italy — Sara Colantonio
Marousi, Greece — Gianna Tsakou
London, UK — Guang Yang

Contents

About the Editors

Ioanna Chouvarda (F), Electrical Engineer, PhD, is an Associate Professor of Medical Informatics and Biomedical Data Analysis, at the School of Medicine, Aristotle University of Thessaloniki. She has been involved in medical informatics research for over 25 years. She is particularly interested in (a) biomedical data analysis and development of machine learning methods with biomedical images, biosignals and biological data, and (b) digital health applications and services incorporating such models. She has worked in numerous national/European research projects in the field of digital health and biomedical informatics, has coordinated Aristotle University of Thessaloniki (AUTH) work in seven projects and has been an evaluator of numerous proposals and projects. She has developed a number of undergraduate and graduate courses in biomedical informatics/engineering, some with an interdisciplinary scope, and has supervised many graduate students. She has been an editorial member of journals and conferences and coordinated the clinical and technical validation AI4HI working group. She is the IEEE Engineering in Medicine and Biology Society (IEEE EMB) Greece chair in 2024. For her published work and citations: https://shorturl.at/pPU49

Sara Colantonio is a senior researcher at the Institute of Information Science and Technologies "Alessandro Faedo" of the National Research Council of Italy, based in Pisa and a member of the "Signals and Images" Laboratory, where she is coordinating the activities of a multidisciplinary team of researchers on the topics of artificial intelligence for health and well-being. Her research interests include artificial intelligence, machine learning, assistive technologies, personal informatics, quantified self, and trustworthy AI. She is co-author of over 80 scientific papers in international journals and conferences (peer-reviewed) and has worked in several national and international projects, dealing with topics related to artificial intelligence applied to the health and well-being of individuals. She coordinated an FP7 project focused on the first sensorized smart mirror for cardiometabolic disease prevention, which granted her an award as one of the 40 Top Transformers in health in 2016. She is a member of the international multi-stakeholder group of experts "Artificial Intelligence for Health Imaging" (AI4HI), contributing to the definition of the FUTURE-AI guidelines on trustworthy AI in medical applications.

Gianna Tsakou is a senior project manager, analyst, and researcher of EC and nationally funded IT research projects. She is currently working as a Senior Project Manager/Senior Analyst in the Research and Development Lab of Maggioli S.P.A., where she has the overview of the eHealth EC-funded projects, notably those related to prevention, diagnosis, and treatment of cancer using big data and AI. Gianna has over 25 years of experience as a researcher and coordinator of EU-funded projects in both the public and private sector. She has participated in more than 30 R&D projects through which she has acquired expertise and know-how in a wide range of IT research areas, including, Artificial Intelligence (AI) in cancer imaging, trustworthy AI, health-related big data repositories, healthy ageing and independent living, eHealth services, accessible services for people with disabilities/ eInclusion, multi-modal interfaces, as well as natural language processing, and machine translation. She has also gained know-how in IT systems for cultural

heritage, digital libraries and eLearning. She is a contributing author of over 30 scientific publications. Since 2003 she is frequently invited by the EC as an evaluator and reviewer of EC projects in a wide variety of scientific areas and under various EC programs. She holds a degree in Greek Philology/Linguistics from the National Kapodistrian University of Athens, having been granted a 4-year scholarship for excellent performance.

Guang Yang is an Associate Professor (Senior Lecturer) in the Bioengineering Department and Imperial-X at Imperial College London. He holds a UKRI Future Leaders Fellowship and serves as an Honorary Senior Lecturer in the School of Biomedical Engineering & Imaging Sciences at King's College London. He is an Associate Editor of *IEEE Transactions* and *npj Digital Medicine*. His research group is dedicated to developing novel and translational techniques for imaging and biomedical data analysis. The group's focus encompasses research and development in data-driven fast imaging, data harmonization, data synthesis, federated learning, explainable AI, and AI in drug discovery. Currently, his work spans a wide range of clinical applications in ageing, cardiovascular disease, lung disease, and oncology. For more information about Yang's Lab, visit: https://www.yanglab.fyi/ and X @gyangMedIA. Email: g.yang@imperial.ac.uk