Trustworthy AI and Data Lineage

Elisa Bertino, *Center for Education and Research in Information Assurance and Security, West Lafayette, IN, 47907, USA*

Suparna Bhattacharya, *Hewlett Packard Labs, Bangalore, 560076, India*

Elena Ferrari, *STRICT Social Lab at the Università degli Studi dell'Insubria, 21100, Varese, Italy*

Dejan Milojicic, *Hewlett Packard Labs, Milpitas, CA, 95035, USA*

AI trustworthiness properties are at the top of concerns for industry, governments, and academia. However, the AI and its models are only as good as the data used to train it. Data lineage could be tracked in many ways, including using metadata, from its generation usage, deployment, and verification. New standards, blueprints, best practices, and repositories for data are required to address requirements for data trustworthiness, such as sustainability, scale, and responsiveness but also ethics, diversity, equity, and inclusion. In this special issue of IEEE Internet Computing*, we feature three articles. The first one addresses certification for trustworthy machine-learning-based applications, the second one is on the topic of data and configuration variances in deep learning, and the third one explores balancing trustworthiness and efficiency in AI Systems. We hope that this special issue will increase the community's awareness of the importance of AI trustworthiness through data lineage.*

We are in the age where AI is deployed and used in almost every facet of our lives and for all products and services we use. However, the AI and its models are only as good as the data used to train it. As the old engineering saying emphasizes, "Garbage in, garbage out." That is the reason that AI trustworthiness properties are at the top of concerns for industry, governments, and academia. This encompasses all phases of the data lifecycle, from generation; to cleansing; to training models, using them, and ensuring that data and models are used appropriately.

This leads us to data lineage, which could be tracked in many ways, including using metadata, from its generation to usage, deployment, and verification. While most phases are ideally automated, some require human engagement, or so-called human-in-the-loop, where ease of use is critical.

As a result, new standards, blueprints, best practices, and repositories for data are required to address functional and nonfunctional requirements for data trustworthiness, such as sustainability, scale, and responsiveness but also ethics, diversity, equity, and inclusion. These are equally important for industry to address the deployment of techniques, policymakers to address governance and regulatory compliance, and academia to advance this important field.

For this special issue, we have received five article submissions, and, after a rigorous review process, we selected two articles and one column.

The article, "Balancing Trustworthiness and Efficiency in Artificial Intelligence Systems: An Analysis of Tradeoffs and Strategies" by Wang[A1] explores the trustworthiness and efficiency of AI systems. The author analyzes the optimal tradeoffs for trustworthy AI. The author specifically explores balances between transparency, robustness, fairness, accountability, efficiency, privacy, and human interaction. The goal is to achieve efficient AI systems that can also have user and societal trust and values.

The article "A Tale of Two Cities: Data and Configuration Variances in Robust Deep Learning" by Zhang et al.[A2] explores the robustness of deep neural network (DNN) models and is motivated by the need for ensuring business and consumer confidence. This article explores DNN robustness as a function of both data

1089-7801 © 2023 IEEE
Digital Object Identifier 10.1109/MIC.2023.3326637
Date of current version 6 November 2023.

APPENDIX: RELATED ARTICLES

A1. Y. Wang, "Balancing trustworthiness and efficiency in artificial intelligence systems: An analysis of tradeoffs and strategies," *IEEE Internet Comput.*, vol. 27, no. 6, pp. 8–12, Nov./Dec. 2023, doi: 10.1109/MIC.2023.3303031.

A2. G. Zhang et al., "A tale of two cities: Data and configuration variances in robust deep learning," *IEEE Internet Comput.*, vol. 27, no. 6, pp. 13–20, Nov./Dec. 2023, doi: 10.1109/MIC.2023.3322283.

A3. M. Anisetti, C. A. Ardagna, N. Bena, and E. Damiani, "Rethinking certification for trustworthy machine learning-based applications," *IEEE Internet Comput.*, vol. 27, no. 6, pp. 22–28, Nov./Dec. 2023, doi: 10.1109/MIC.2023.3322327.

and software variability. The authors use search-based optimization to devise robust learning in the presence of representative variances in data and software configurations.

Finally "Rethinking Certification for Trustworthy Machine Learning-Based Applications" by Anisetti et al.[A3] addresses how machine learning (ML) is applied to advanced applications with nondeterministic behavior. The authors specifically look at the application that operates on the cloud–edge continuum. They advocate for assurance solutions for nonfunctional properties to improve their trustworthiness. Certification is of high importance to industry and governments as an assurance policy and regulation technique for data trustworthiness. The authors analyze current certification schemes and propose an ML-based certification scheme.

We hope that this special issue will increase the community's awareness of the importance of AI trustworthiness through data lineage. We hope that our community will learn as much as we did in the process of putting together this special issue. We would like to thank the authors of all accepted and submitted articles as well as the reviewers who helped us in the selection process. It takes a (community) village to build trust.

ELISA BERTINO is a professor of computer science at Purdue University and is acting as the research director of the Center for Education and Research in Information Assurance and Security, West Lafayette, IN, 47907, USA. Contact her at bertino@purdue.edu.

SUPARNA BHATTACHARYA is an HPE Fellow and vice president at Hewlett Packard Labs, Bangalore, 560076, India. Contact her at suparna.bhattacharya@hpe.com.

ELENA FERRARI is a professor of computer science and director of the STRICT (Security and TRust for Information and Communication Technology) Social Lab at the Università degli Studi dell'Insubria, 21100, Varese, Italy. Contact her at elena.ferrari@uninsubria.it.

DEJAN MILOJICIC is an HPE Fellow and vice president at Hewlett Packard Labs, Milpitas, CA, 95035, USA. Contact him at dejan.milojicic@hpe.com.