Protocol for Systematic Literature Review in Trustworthy AI supported by blockchain

Oussema Abdelmoula
CRI, Université Paris 1 Panthéon-Sorbonne
Université Paris 1 Panthéon-Sorbonne
Paris, France
oussema.abdelmoula@gmail.com

***Abstract*—Enhancing the integrity and transparency of AI systems is made possible by the combination of blockchain technology with artificial intelligence (AI). To find out how blockchain may help reliable AI, this research performs a systematic literature review (SLR). Our goal is to demonstrate how blockchain technology can be used to guarantee data integrity, increase the accountability of machine learning processes, and make AI judgments more explainable by examining previous research. This analysis highlights areas for further investigation and suggests approaches for creating trustworthy, moral, and open AI systems with blockchain technology.**

***Index Terms*—blockchain, artificial intelligence, trustworthy AI, machine learning, data management, inference**

I. Motivation

Examining the connection between blockchain innovation and artificial intelligence is the main thrust behind this precise writing survey (SLR), with an exceptional accentuation on how blockchain can support trustworthy artificial intelligence. Trustworthy AI encompasses several key aspects: trustworthy machine learning, trustworthy data management, and trustworthy inference [1]. Since blockchain innovation is transparent and changeless, it offers a promising way to enhance trust in AI systems. This SLR seeks to summarize previous research, point out gaps, and suggest future research directions in leveraging blockchain for trustworthy AI [1], [2], and [3].

II. Research Questions

The review will address the following research questions (RQs):

- **RQ1:** How can blockchain be utilized to build a trustworthy data preparation pipeline for AI?
- **RQ2:** How can blockchain support trustworthy machine learning?
- **RQ3:** How can blockchain assure the explainability of AI inference?
- **RQ4:** How can trustworthy AI components be assembled to build a fully trustworthy end-to-end AI solution?

III. Search Query

To ensure a thorough investigation of relevant material, we will utilize the Scopus metasearch engine using the following search query:

((blockchain) AND (”artificial intelligence” OR ”AI”) AND (trust* OR ”data integrity” OR ”explainable AI” OR ”Federated Learning” OR ”privacy-preserving” OR ”data provenance”))

IV. Inclusion/Exclusion Criteria

We will use the following inclusion and exclusion criteria to choose pertinent studies:

A. Inclusion Criteria

- Papers presented at conferences or in peer-reviewed journals are the first requirement for inclusion.
- Research concentrating on using blockchain technology to improve AI's credibility.
- English-language articles.
- Works published after 2010 that document new developments.

B. Exclusion Criteria

- Non-peer-reviewed articles, such as opinion pieces, editorials, and white papers, are excluded based on the following criteria.
- Research not specifically pertaining to the trustworthy intersection of blockchain and AI.
- Studies that are duplicates from various databases.

V. Review Protocol

The methodical procedure to be followed during the SLR is described in the review protocol [4]. It makes sure the review process is thorough, transparent, and replicable [5].

A. Preparing the Review

- Establish research questions in light of the study's goals and motivation.
- Create a thorough search query that is specific to the study questions.
- Determine which databases and meta-search engines to use for the literature search (this will be Scopus for our study).

B. Review Procedure

- **Search Strategy:** Run the query through the chosen databases.
- **Concentrate on Determination:** Use incorporation and avoidance guidelines to limit relevant examination. This procedure includes:
  - **Initial Screening:** Examine abstracts and titles to weed out unrelated research.
  - **Full-text Review:** Evaluate the remaining studies in their entirety to make a final inclusion decision.
  - **Data Extraction:** Gather data from chosen studies utilizing a standardized form for data extraction, making sure to record important details like:
    * Concentrate on metadata (title, creators, year of distribution, source)
    * Research objectives
    * Methods applied
    * Comes about approximately the unwavering quality of blockchain and AI
    * Gaps identified and recommendations for further investigation

C. Presenting the Review

- **Synthesis of Findings:** Condense and integrate the data that was taken out to respond to the research questions.
- **Presentation:** Arrange the data into a well-structured report that includes parts about the approach, results, discussion, and conclusion.
- **Dissemination:** Present the findings at pertinent conferences and in scholarly journals to spread the word about the findings.

VI. Results

The reporting of this SLR's findings shall follow the instructions given by Kitchenham et al. (2009) [4]. A thorough summary of the body of research will be included in the paper, emphasizing how in which blockchain technology may facilitate reliable AI. It will point out areas in need of research and suggest avenues for further investigation, advancing the creation of reliable, open, and moral AI systems [6].

VII. References

References

[1] E. Toreini, M. Aitken, K. Coopamootoo, K. Elliott, C. G. Zelaya, and A. van Moorsel, "The relationship between trust in AI and trustworthy machine learning technologies," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ser. FAT* '20. Association for Computing Machinery, pp. 272–283.

[2] M. Nassar, K. Salah, M. H. ur Rehman, and D. Svetinovic, "Blockchain for explainable and trustworthy artificial intelligence," vol. 10, no. 1, p. e1340, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/widm.1340.

[3] V. Tolpegin, S. Truex, M. E. Gursoy, and L. Liu, "Data poisoning attacks against federated learning systems."

[4] B. Kitchenham, P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering-a systematic literature review," vol. 51, pp. 7–15.

[5] N. Six, A. Perrichon-Chrétien, and N. Herbaut, "SAIaaS: A blockchain-based solution for secure artificial intelligence as-a-service," in *The 2nd International Conference on Deep Learning, Big Data and Blockchain (DEEP-BDB 2021)*.

[6] W. Liang, G. A. Tadesse, D. Ho, F.-F. Li, M. Zaharia, C. Zhang, and J. Zou, "Advances, challenges and opportunities in creating data for trustworthy AI."