

Knowledge Representation for Algorithmic Auditing to Detangle Systemic Bias

Paola DiMaio

Center for Systems, Knowledge Representation and Neuroscience

Abstract: Knowledge is central to cognition, and adequate knowledge representation is necessary to ensure the correct functioning of intelligent systems leveraging knowledge. Knowledge auditing methodologies in use, however do not typically address the requirements for evaluating the consistency of logical and functional knowledge representation constructs, as used in AI, that support the correct execution of system processes. This paper puts forward the notion that algorithmic bias results from the lack of adequate knowledge representation mechanisms, and is part of systemic bias. The notion of systemic bias is identified, characterized, and described as an emergent relation resulting from the combination of other factors between knowledge representation (KR) and bias in AI. These factors are here named conceptualised, related to each other and visualised. The resulting artefacts are then incorporated in the KAF (Knowledge Audit Framework) methodology (first published in 2012) and adopted as a reference model for understanding and representing bias as a systemic emergent phenomenon.

Table of Contents

- | | |
|--|--|
| 1. Introduction and Motivation | |
| 2. Semantics and Knowledge Representation | |
| 3. Accountability And Algorithmic Bias | |
| 3.1. Disentangling Systemic Bias | |
| 3.2. Knowledge Representation | |
| 3.3. Deep fakes As Knowledge Misrepresentation | |
| 4. Auditing: Data, Knowledge, Algorithms | |
| 4.1. Data Auditing | |
| 4.2. Knowledge Auditing | |
| 4.3. The Knowledge Audit Framework | |
| 4.4. Algorithm Auditing | |
| | 5. Knowledge Representation For Algorithmic Audits |
| | 5.1. Criteria for Auditing KR |
| | 5.2. Levels of Representation |
| | 5.3. Adequacy |
| | 5.4 Algorithmic Auditability |
| | 6. Adding KR To KAF |
| | 7. Conclusions and Future Work |

The full research paper is available as an open access resource via Open access repositories or by contacting the author.

Paola.dimaio@gmail.com

ISIC'21: International Semantic Intelligence Conference, February 25–27, 2021, New Delhi, India

✉ : paola.dimaio@gmail.com (P.D. Maio)



Copyright © 2021 for this paper by the authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)