SemeAd 2023
XXVI Seminários em Administração

07, 08, 09 e 10 de novembro de 2023
ISSN 2177-3866

SUSTAINABLE PROJECT MANAGEMENT FOR THE FOURTH INDUSTRIAL REVOLUTION: SYSTEMATIC REVIEW ON DATA QUALITY EVALUATION OF DATA USED IN COGNITIVE COMPUTING AND AI

MARCO AURÉLIO MAZZEI
UNIVERSIDADE PAULISTA (UNIP)

ARNALDO LUIZ RYNGELBLUM
UNIVERSIDADE PAULISTA (UNIP)

MAURO LUIZ MARTENS
UNIVERSIDADE PAULISTA (UNIP)

Agradecimento à orgão de fomento:
This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

SemeAd 2023
XXVI Seminários em Administração

07, 08, 09 e 10 de novembro de 2023
ISSN 2177-3866

SUSTAINABLE PROJECT MANAGEMENT FOR THE FOURTH INDUSTRIAL REVOLUTION: SYSTEMATIC REVIEW ON DATA QUALITY EVALUATION OF DATA USED IN COGNITIVE COMPUTING AND AI

Introdução

Effective use of AI tools depends on our understanding of how the quality and representativeness of data used in AI training and Cognitive Computing can impact algorithm development and how reliable they can be. A systematic review of literature was done to understand how the propositions to use AI to support sustainability in project management evaluates data quality in its technological, linguistic, and sociocultural dimensions and in a new front, propose implementation of guards against improper feedback generated synthetically by Larga Language Models.

Problema de Pesquisa e Objetivo

When proposing the use of AI to support sustainability in project management, attention should be paid to the quality of data and feedback used in algorithm training. Projects that propose such use must work out how to obtain data quality in all its dimensions. Understand how research and the production of academic knowledge discuss and evaluate the quality and representativeness of the data used in the training of AI algorithms and if they are considering strategies to improve their quality and correct human training to train algorithms.

Fundamentação Teórica

To assess data quality considerations on the articles, we were able to establish a framework based on ongoing discussions by grouping the types of problem sources into four categories: raw data, synthetic data generated, linguistic, and sociocultural type errors. Such a framework was developed considering articles in the areas of artificial intelligence and cognitive computing development that are already facing the results of these types of errors and alerting to the importance of training to improve technology results.

Discussão

The academic community already understands raw data quality as a prerequisite for Industry 4.0 technologies and explicitly mentions it in all fourteen articles. As for the discussion of linguistic aspects of data quality, eight articles list this as an important factor. Of the set of articles, only four discussed the sociocultural aspects of data collection and use and none discussed the issue of synthetic data generated and its consequences with the popularization of large language models leading generative artificial intelligence to the masses

Conclusão

There is a risk in automating decision-making processes based on algorithms trained with poor data quality or with a lack of sample representativeness. This risk is already on the research radar However, there is also a new risk arising from using textually biased data to train algorithms that will absorb this bias and hide bad decisions or make them invisible. With the advancement of technology, new risks arise with the use of the Large Language Model tools to do the training replacing the human work in giving feedback on the results of the algorithm being trained.

Referências Bibliográficas

Aljawder, A., & Al-Karaghouli, W. (2022). The adoption of technology management principles and artificial intelligence for a sustainable lean construction industry in the case of Bahrain. Journal of Decision Systems, 1–30. https://doi.org/10.1080/12460125.2022.2075529 Aria, M., & Cuccurullo, C.

07, 08, 09 e 10 de novembro de 2023
ISSN 2177-3866

(2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007 Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic researc