Research Article

Topic Modelling and Sentiment Analysis of AI-driven Circular Procurement practices

Temitope Omotayo[1], Abisola Muftau[1], Jiamei Deng[1], Ali Murat Tanyer[2], Nilüfer Kızılkaya Öksüz[3], Anderson Akponeware[1]

1 School of Built Environment, Engineering and Computing. Leeds Beckett University, Leeds
2 Research Center for the Built Environment (Research Center for the Built Environment (MATPUM), Middle East Technical University, Ankara, & Turkey).
3 Department of Architecture, Building Science Program. Middle East Technical University, Ankara, Turkey.
Correspondence: a.o.muftau@leedsbeckett.ac.uk

Abstract

The construction industry is the biggest worldwide consumer of raw materials, producing the most waste among all industries, with the UK construction sector generating 62% of national waste. Organisations can address these challenges through Circular Procurement (CP), which implements Circular Economy (CE) principles by incorporating reuse and recycling and material efficiency into procurement processes. However, CP remains underexplored in construction scholarship and practice. The research uses sentiment analysis and topic modelling to explore the integration of Artificial Intelligence (AI) in CP discourse. The research included 23 peer-reviewed articles from Scopus and Google Scholar databases. The sentiment analysis showed a positive attitude toward AI procurement, but most of the content was neutral because it focused on description rather than emotional expression. The LDA model produced five thematic areas, including machine learning and circularity. The corpus shows that circular procurement appears in only 3% of the text, which indicates its early stage of development in this field. The results show that AI methods already simplify procurement and could be extended to embed CE principles more effectively in construction. The originality of this research lies in combining sentiment analysis with topic modelling to provide a dual perspective on both tone and latent themes of the academic discourse. The research value comes from its demonstration of limited academic studies on AI-driven CP and its presentation of how existing AI procurement tools can be modified to support circularity. The research addresses current knowledge deficiencies while creating fundamental elements for sustainable procurement methods in construction projects.

Keywords: Artificial Intelligence; Circular Procurement; Construction; Sentiment Analysis; Topic Modelling

Highlights

- AI-driven procurement shows strong potential for embedding circular economy principles.
- Circular procurement remains nascent, with limited but growing academic attention.
- There is a predominantly positive outlook on the use of AI in construction procurement.

1 Introduction

As the construction industry accounts for the highest consumption of resources and materials of all the different industries, so does waste generated. The UK construction industry accounts for 62% of waste in comparison to other industries (Ali, Ali, & Bayyati, 2018). Understanding the sources is important to proffer effective waste management solutions. Construction and demolition (C&D) waste constitutes over half of UK's landfill, similarly in Turkey, where it is the largest volume of solid waste being produced (Arslan, Coşgun, & Salgın, 2012). In their quantitative analysis of sources of construction waste using a questionnaire, Fadiya et al. (2014) sampled a representation of building and civil engineering contractors operating locally and nationally in the UK. The research categorized the sources of waste into procurement, design, material handling, operations, weather, vandalism, misplacement, residual, and others. The study alluded that a reliable evaluation of construction waste prior to the start of construction activities is critical to understanding financial implications of waste generation, and aid decision making for waste reduction. Ali et al. (2018) similarly corroborate the issue of waste generation in the UK construction industry, accentuating factors contributing to waste generation, and noting how sustainability can be achieved through waste reduction. The current construction sector is heavily reliant on the linear model of consumption and disposing of products and materials. This is in contrast with circular economy (CE) principles of reducing construction waste from the inception to completion phases of projects through the reuse, recycling, minimization, and recovery of high-value construction materials (Ellen Macarthur Foundation, 2020; Korhonen, Honkasalo, & Seppälä, 2018). Circular procurement (CP) is a nascent concept by which CE principles can be achieved in the construction industry (Sönnichsen & Clement, 2020). Circularity (Antwi-Afari, Ng, & Hossain, 2021) involves the use of high-value reusable and recyclable materials in the design of products. A circular product can be defined in terms of its degree of purchase, the most circular product being the one not purchased at all, while the least circular product is the one that needs to be purchased. In a CE, circular products are used circularly. With efforts geared towards a CE and reducing material consumption and waste, world economies are transitioning and transforming towards more sustainable, low-carbon, and resource-efficient practices, with the European Union (EU) leading the pack (European Commission, 2015).

The principles of a CE in the construction landscape focus on improving the use of sustainable materials, as opined by Akanbi et al. (2018) and Herczeg et al. (2018); promoting materials efficiency through the reuse and recycling of construction materials (Ghisellini, Ripa, & Ulgiati, 2018); and avoiding unnecessary waste generation (Hossain & Ng, 2018). Guerra et al. (2021) presented empirical evidence detailing the level of implementation of circular economy principles in the construction industry. The research developed a knowledge framework, identified and selected construction companies, and collected data, with the aim of identifying trends, gaps, and providing recommendations. The study went on to reveal that most of the businesses spearheading the transition towards CE were newer and smaller sized. Some principles of CE such as Waste as a Resource, Resource Recovery, and Circular Supplies were easier and faster to adopt. At the same time, complex circular strategies like Sharing Platform and Product as a Service (PaaS) were more difficult to adopt. Antwi-Afari et al. (2021) went further by conducting a scientometric analysis to evaluate circularity gaps in the construction industry. The findings identified key themes of CE in the construction industry, circularity gaps, and suggestions to bridge the gaps.

CP is the mechanism through which the principles of circularity can be achieved within the construction industry. Fuertes Giné et al. (2022) posit that CP involves the purchase of goods and services, to promote CE practices of reuse, recycle and waste minimization. Few studies have been conducted with regard to the adoption of CP across different sectors of the economy, such as furniture (Ntsonde & Aggeri, 2021), food (Ada et al., 2021), beverages (Urbinati, Chiaroni, & Toletti, 2019), etc. Similarly, limited research is also available with respect to the construction industry. Coenen (2019) developed a framework as an assessment method in the design and procurement process for the construction of circular bridges and viaducts. This entails evaluating and comparing different design options for construction projects, allowing the incorporation of circularity as a selection metric within the procurement process, thereby benefiting both designers at the market side and client project managers. Ultimately, project managers are able to determine how circularity affects the awarding of contracts. The study notes that since CP rewards the degree of circularity of a business, construction projects should be designed in a way as to be able to demonstrate their circularity through indicators -design input, adaptability, resource availability, reusability -that are measurable. This will encourage contractors to create designs that are rated high on specified parameters. Furthermore, Sajid et al. (2024) highlight the advantages CP, focused on sustainability, has over traditional procurement, which is a linear model. The findings explain the role CP plays in resource optimization, emission reduction, energy conservation and social sustainability in the construction industry. It further notes that although several literatures exist on barriers to adopting CE in the built environment, there is still a dearth of comprehensive reviews on barriers to adopting CP in the construction industry. It, however, highlights that Europe is the highest regional contributor to CP research, with the UK leading. The research finally analyses the key barriers to CP adoption in CI, categorizing them into hard and soft barriers, listing specific strategies for addressing them, and presenting a conceptual framework for promoting CP in the built environment, one of which is exploring the role of new and emerging technologies.

Text mining has shown usefulness in extracting insights from textual data (Tabassum & Patil, 2020), however, there is no study applying these techniques to investigate AI-driven circular procurement practices. The objectives of this paper is (i) To identify the evolving of AI-driven practices in CP; (ii) To explore recent trends in the adoption of AI-driven technologies in CP; and (iii) To suggest future research opportunities for AI-driven technologies in CP.

2 Methodology

This study conducted a literature review of existing research. It employed a qualitative approach (Tisdell, Merriam, & Stuckey-Peyrot, 2025) using secondary data in order to find answers to complex research problems. This research used data from Scopus database and Google Scholar. A direct search was conducted on Scopus database, and Harzing's Publish or Perish (Harzing A.W, 2007) was used as the search tool to extract articles from Google Scholar. One hundred and sixty-eight (165) documents were retrieved from Scopus, while Google scholar search was limited to two hundred (200) results. The scope of the literature was limited by considering on English articles between 2000 and 2024. Books and thesis were excluded; the remaining documents were manually reviewed further to eliminate irrelevant articles and EndNote referencing tool was used to identify duplicate documents. The documents were merged, each read in-depth, and a total of twenty-three (23) peer-reviewed pdf articles were considered appropriate to be included in the analysis. Sentiment analysis and topic modelling, specifically Latent Dirichlet Allocation (LDA) were used to assess opinion and identify latent themes from the full-text data. The combination of both techniques helps to uncover latent topics in unstructured data and reveal the emotions surrounding the text (Qiao & Williams, 2022; Tleuken, Orel, Iskakova, Varol, & Karaca, 2024). The comprehensive flowchart for this study is presented in Figure 1.

2.1 Sentiment Analysis

Analyzing unlabelled data using sentiment analysis allows for the understanding of large amounts of text to quickly get a sense of the sentiment expressed. For this study, we employ a pre-trained

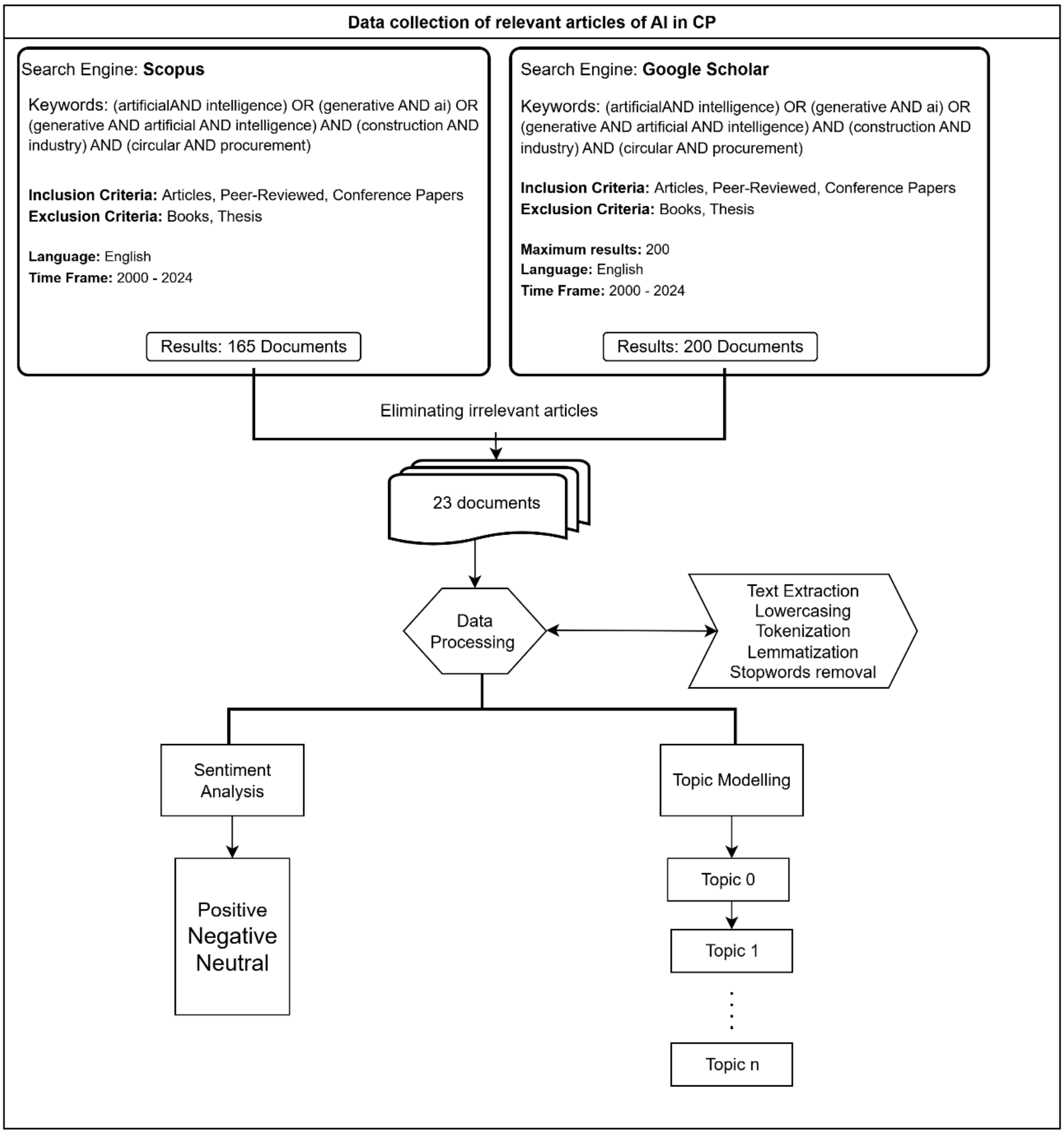

Figure 1. Review Flowchart

sentiment analysis model to predict sentiments for each piece of unlabelled text. The transformer model used is ***distilbert-base-uncased-finetuned-sst-2-english***, a pre-trained model from the Hugging Face Transformers library. This model is a faster, lighter model of the original BERT (Bidirectional Encoder representations from Transformers), with a standard base size architecture trained on lowercased text, which has been fine-tuned specifically for sentiment analysis on the SST-2 (Stanford Sentiment Treebank v2). The data processing involves extracting raw text from the corpus. A document-level and paragraph-level segmentation was conducted on the extracted text, and the preprocessed document and paragraph texts are fed in batches into the sentiment model pipeline to predict the sentiment labels. Finally, insightful visual representations of the analysis are generated through document and paragraph sentiment distribution plots.

2.2 Topic modelling

The steps involved in topic modelling using Latent Dirichlet Allocation include: (i) extracting text from the PDF files using pdfminer, (ii) preprocessing the text by converting all text to lowercase, cleaning, tokenization, and lemmatization, (iii) building a dictionary from the processed tokens and creating a corpus by converting the processed tokens into a Bag-of-Words representation, (iv) train the model using the corpus and dictionary using the Gensim Python library for topic modelling, and finally (v) interpret the topics and visualise topics with pyLDAvis.

3 Results

The map in Figure 2 shows a key concentration of documents in the UK, Europe, North America, and parts of Asia. The UK is the region with the most prominent source of documents, followed by the USA and Germany. There is a geographical imbalance in the origin of journal articles relating to procurement, circular economy and AI in construction, in which there is the dominance of Western countries, with Africa, Russia, and Southeast Asia having very few data points. There are no data points in South America.

Figure 2. Map of document origin

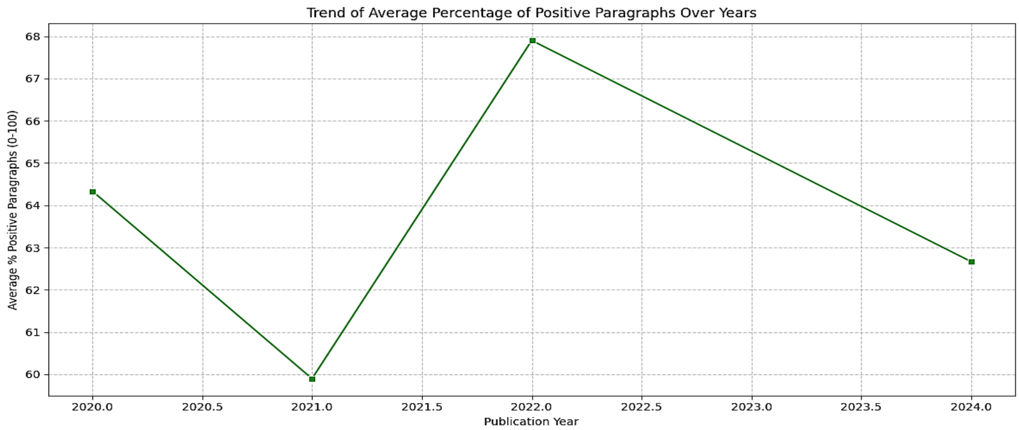

Figure 3. Sentiment trend according to publication year

Figure 3 shows the changes in sentiments over the publication year period, showing trends for 2020 till 2024. The average percentage of positive paragraphs related to AI techniques in procurement in 2020 stood at approximately 64%, with a sharp decline to 60% by 2021. After this came a period of steep increase in 2022, rising as high as approximately 68%. This accounts for the most positive year in the observed publication period. Although there is a subsequent decline by 2024, the level remained higher than the 2021 value.

3.1 Sentiment analysis

The chart in Figure 4 shows the categorisation of the documents based on their majority sentiment labels, classified as either "Positive", "Negative" or "Neutral". The overall tone of the analysis is predominantly positive, accounting for around 20 documents, with 3 documents labelled as negative, revealing that all documents have a clear categorisation. A closer look at the distribution of the paragraphs across different sentiment label classifications shows that the most common sentiment label is neutral, accounting for approximately 13,000 paragraphs, likely due to the content of the paragraph not conveying a strong sentiment. This suggests that most paragraphs convey factual information rather than emotive content. There is also a representation of positive and negative paragraph labels indicating the varying topics in the text.

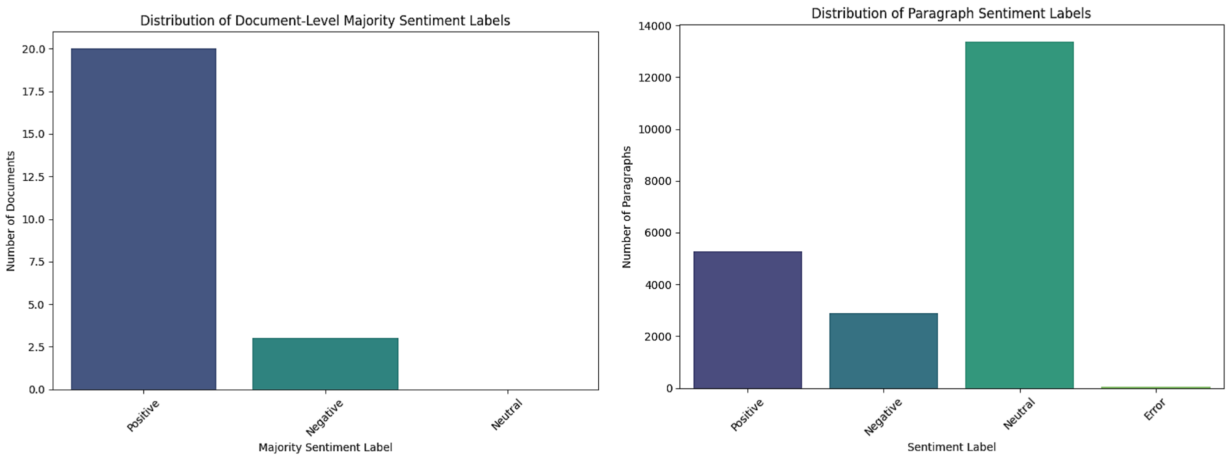

Figure 4. Sentiment Labels across document and paragraph

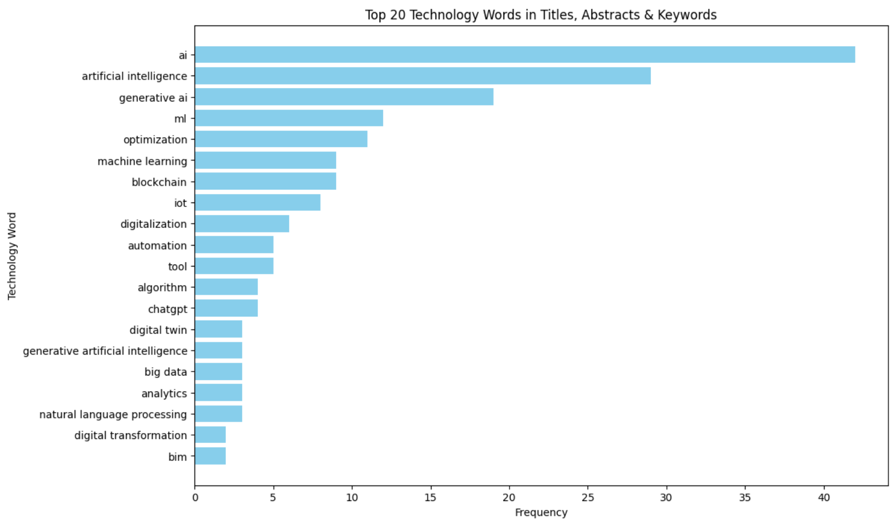

Figure 5. Technology-Related Words

Figure 5 shows the analysis of technology-related terms in the corpus' titles, abstracts and keywords. It reveals varying technologies that are being applied in the context of the construction industry aimed at improving procurement processes. The integration of AI into procurement practices is evident by it accounting for the most common technology being integrated into procurement, followed by other technologies like generative AI, machine learning, big data and digital twin; such as the use of big data in forecasting and minimize possible inefficiencies throughout a construction lifecycle in CP (Awan, Sroufe, & Shahbaz, 2021).

3.2 Topic Modelling

In analyzing the corpus using LDA, coherence score was used in determining the optimum number of topics (Syed & Spruit, 2017), which was five. Topic 2 constituted the popular theme accounting for 72% of the tokens, followed by Topic 3 with 14.2%, and Topic 3 being the smallest with 3%. Figure 5 shows the top 10 words in each topic. The top 5 words generated from Topic 0 was "ml", "cp", "csc", "waste", "product"; "uncertainty", "model", "supply", "demand", "price" were generated for Topic 1, while Topic 2 words were "ai", "construction", "datum", "procurement", "model". Topic 3 and Topic 4 were "model", "cost", "cid", "fuzzy", "method"; and "procurement", "blockchain", "technology", "public", "sector" respectively.

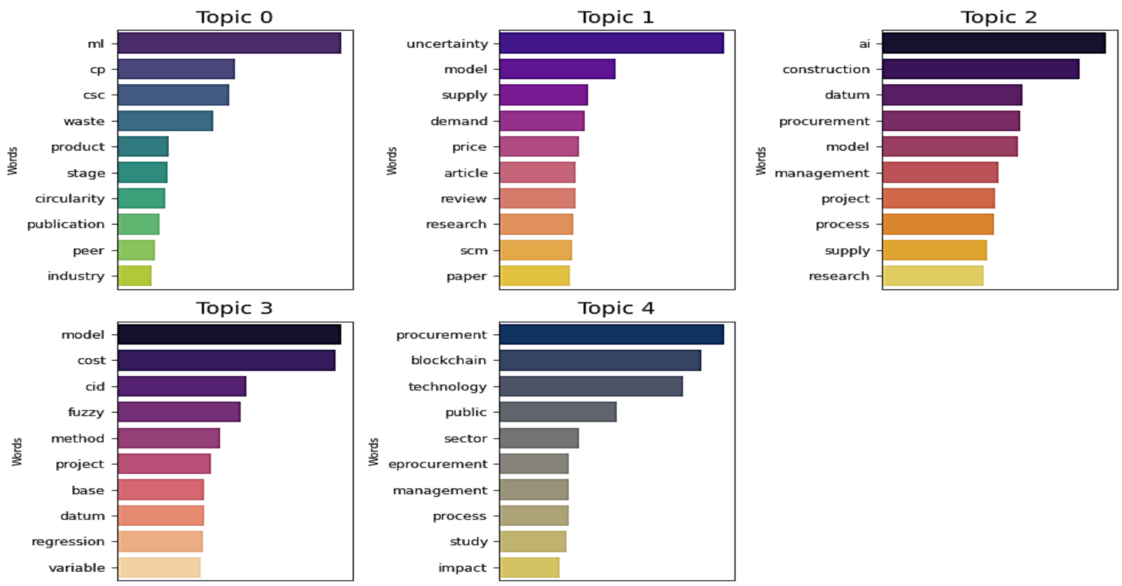

Figure 5. LDA Topics.

The document heatmap in Figure 6 also shows the association of each document with the extracted topics. The vertical rows list the documents, with the distribution of topics in each document represented by the horizontal bars. Values 0.0 – 1.0 depict the probability of the document's content to the topic, while the colour coding revealing the dominant topics for each document; showing whether a document is single-topic inclined such as "Application of Artificial Intelligence (AI) in Sustainable Building Lifecycle" or a multi-topic mix like "Artificial Intelligence (AI) in E-Procurement", with a lot of the documents being clustered around Topic 4.

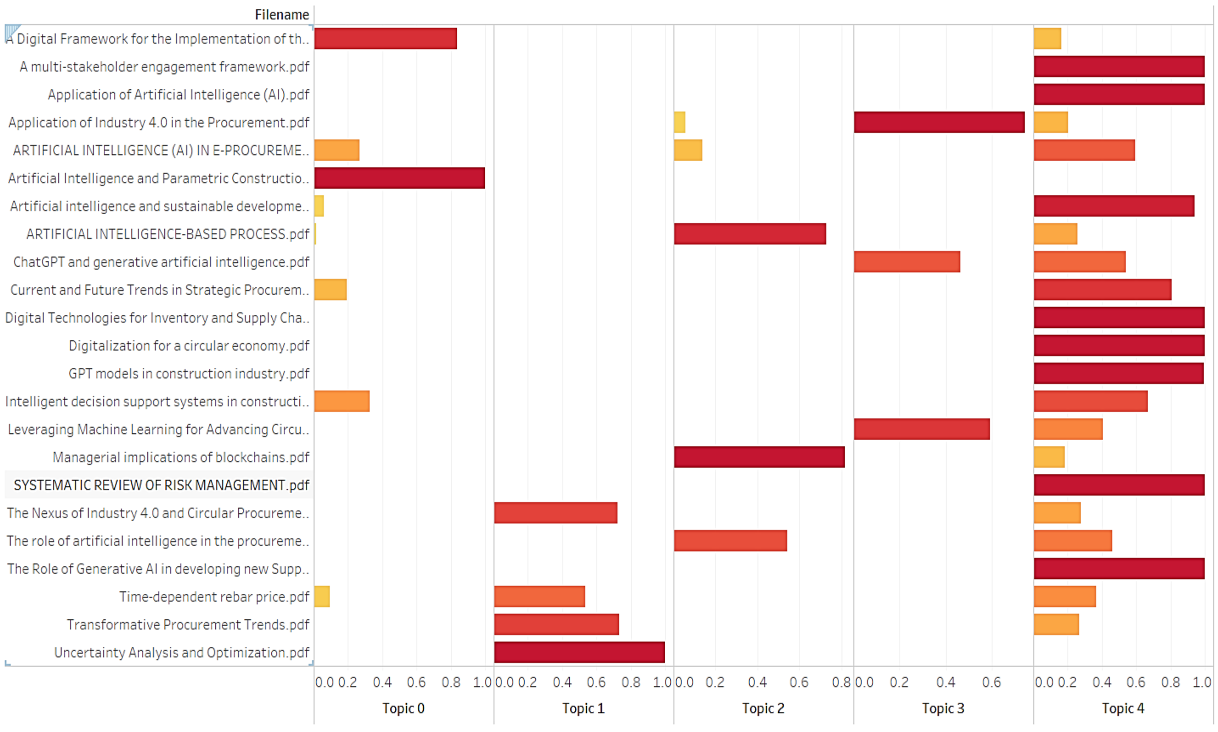

Figure 6. Document Heatmap.

4 Discussion

Analysis reveal that the corpus of text examines the use of digital technologies in supply chain economics, procurement, and achieving sustainability in construction. The topics that research into CP is about 3% of the entire corpus, further affirming the nascent nature of this concept. Digital technologies are helping to enable circularity principles, with a lot of the articles emphasizing the role of technologies such as AI, IoT, ML, DL, BIM, in the optimization of resources, promoting material reuse, extending the life cycle of products, enabling predictive maintenance, enhancing data-driven analysis, and cost modelling. The role of AI-powered systems in waste management, which is key to reuse, reduce, recycle was also identified in some texts. This research also reveals strides achieved in incorporating AI into general procurement practices (Li, Culmone, De Reyck, & Yoo, 2025) , shown by the larger percentage of positive sentiment.

Themes generated from topic modelling include Machine learning and circularity, Uncertainty in market dynamics, AI in construction procurement, Cost modelling techniques, and Advancing public procurement. These themes agree with the sentiment analysis, indicating the applications of ML and AI in CP in the construction industry. It also reveals some other research work being developed to achieve sustainability in procurement processes through the development of methods such as green public procurement, circular public procurement (Kristensen, Mosgaard, & Remmen, 2021).

Although a thorough search was conducted, this study was limited by the search criteria for the selected databases. In all likelihood, there is the existence of other articles on the subject, but they were not found. The potential bias in the manual selection of final papers and geographical imbalance further limits the discourse to the only available regions.

5 Conclusions

Employing AI-driven technologies in achieving circular procurement in construction is in its infancy stages but gaining traction. This research spotlight the relevance of circular procurement practices in construction, and efforts the industry is taking in achieving circularity by leveraging AI technologies. The prevalence of positive sentiments in AI-driven procurement shows the trends in adopting this innovation and the topic modelling themes shows that research into CP in construction and the integration of AI is largely underexplored. The evident positive sentiment towards AI-driven solutions in geographical regions identified by this study could be worth exploring by other regions. The methodology of this study based on secondary data could be expanded by applying a mixed method approach, or by investigating and analyzing case studies from leading circular economies.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgements

The authors are grateful for the support provided by the International Science Partnership Fund (ISPF)

Funding

This project is funded by the International Science Partnership Fund (ISPF) Grant Number: 1203563630.

Data Availability Statement

All data are available upon request.

References

Journal Article
Ada, N., Kazancoglu, Y., Sezer, M. D., Ede-Senturk, C., Ozer, I., & Ram, M. (2021). Analyzing barriers of circular food supply chains and proposing industry 4.0 solutions. *Sustainability, 13*(12), 6812.
Journal Article
Akanbi, L. A., Oyedele, L. O., Akinade, O. O., Ajayi, A. O., Delgado, M. D., Bilal, M., & Bello, S. A. (2018). Salvaging building materials in a circular economy: A BIM-based whole-life performance estimator. *Resources, Conservation and Recycling, 129*, 175-186.
Conference Paper
Ali, S., Ali, A., & Bayyati, A. (2018). *Achieving Sustainability in the UK Construction by Reducing Waste Generation*. Paper presented at the PSYCON3: Psycon-'Psychology & Sustainable Construction': Searching the gap between psychology and construction for a sustainable built environment.
Journal Article
Antwi-Afari, P., Ng, S. T., & Hossain, M. U. (2021). A review of the circularity gap in the construction industry through scientometric analysis. *Journal of cleaner production, 298*, 126870.
Book
Arslan, H., Coşgun, N., & Salgın, B. (2012). *Waste Management in Turkey*.
Journal Article
Awan, U., Sroufe, R., & Shahbaz, M. (2021). Industry 4.0 and the circular economy: A literature review and recommendations for future research. *Business Strategy and the Environment, 30*(4), 2038-2060.
Journal Article
Coenen, T. B. (2019). Circular bridges and viaducts: Development of a circularity assessment framework.
Electronic Article
Ellen Macarthur Foundation. (2020). What is a Circular Economy. Retrieved from https://www.ellenmacarthurfoundation.org/topics/circular-economy-introduction
Government Document
European Commission. (2015). *Closing the loop - An EU action plan for the Circular Economy*. Brussels Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52015DC0614
Journal Article
Fadiya, O. O., Georgakis, P., & Chinyio, E. (2014). Quantitative analysis of the sources of construction waste. *Journal of Construction Engineering, 2014*(1), 651060.
Journal Article
Fuertes Giné, L., Vanacore, E., & Hunka, A. D. (2022). Public procurement for the circular economy: A comparative study of Sweden and Spain. *Circular Economy and Sustainability, 2*(3), 1021-1041.
Journal Article

Ghisellini, P., Ripa, M., & Ulgiati, S. (2018). Exploring environmental and economic costs and benefits of a circular economy approach to the construction and demolition sector. A literature review. *Journal of cleaner production, 178*, 618-643.

Journal Article

Guerra, B. C., Shahi, S., Mollaei, A., Skaf, N., Weber, O., Leite, F., & Haas, C. (2021). Circular economy applications in the construction industry: A global scan of trends and opportunities. *Journal of cleaner production, 324*, 129125.

Computer Program

Harzing A.W. (2007). Publish or Perish. Retrieved from https://harzing.com/resources/publish-or-perish

Journal Article

Herczeg, G., Akkerman, R., & Hauschild, M. Z. (2018). Supply chain collaboration in industrial symbiosis networks. *Journal of cleaner production, 171*, 1058-1067.

Journal Article

Hossain, M. U., & Ng, S. T. (2018). Critical consideration of buildings' environmental impact assessment towards adoption of circular economy: An analytical review. *Journal of cleaner production, 205*, 763-780.

Journal Article

Korhonen, J., Honkasalo, A., & Seppälä, J. (2018). Circular economy: the concept and its limitations. *Ecological economics, 143*, 37-46.

Journal Article

Kristensen, H. S., Mosgaard, M. A., & Remmen, A. (2021). Circular public procurement practices in Danish municipalities. *Journal of cleaner production, 281*, 124962.

Journal Article

Li, X., Culmone, V., De Reyck, B., & Yoo, O. S. (2025). Automating Procurement Practices Using Artificial Intelligence. *INFORMS Journal on Applied Analytics, 55*(3), 195-223.

Journal Article

Ntsonde, J., & Aggeri, F. (2021). Stimulating innovation and creating new markets–The potential of circular public procurement. *Journal of cleaner production, 308*, 127303.

Journal Article

Qiao, F., & Williams, J. (2022). Topic modelling and sentiment analysis of global warming tweets: evidence from big data analysis. *Journal of Organizational and End User Computing (JOEUC), 34*(3), 1-18.

Unpublished Work

Research Center for the Built Environment (MATPUM), Middle East Technical University, Ankara, & Turkey.

Journal Article

Sajid, Z. W., Aftab, U., & Ullah, F. (2024). Barriers to adopting circular procurement in the construction industry: The way forward. *Sustainable Futures, 8*, 100244.

Journal Article

Sönnichsen, S. D., & Clement, J. (2020). Review of green and sustainable public procurement: Towards circular public procurement. *Journal of cleaner production, 245*, 118901.

Conference Proceedings

Syed, S., & Spruit, M. (2017). *Full-text or abstract? examining topic coherence scores using latent dirichlet allocation*. Paper presented at the 2017 IEEE International conference on data science and advanced analytics (DSAA).

Journal Article

Tabassum, A., & Patil, R. R. (2020). A survey on text pre-processing & feature extraction techniques in natural language processing. *International Research Journal of Engineering and Technology (IRJET), 7*(06), 4864-4867.

Book

Tisdell, E. J., Merriam, S. B., & Stuckey-Peyrot, H. L. (2025). *Qualitative research: A guide to design and implementation*: John Wiley & Sons.

Journal Article

Tleuken, A., Orel, D., Iskakova, A., Varol, H. A., & Karaca, F. (2024). Exploring stakeholders' opinions on circular economy in the construction sector: A natural language processing analysis of social media discourse. *Journal of Industrial Ecology, 28*(4), 853-867.

Journal Article

Urbinati, A., Chiaroni, D., & Toletti, G. (2019). Managing the introduction of circular products: Evidence from the beverage industry. *Sustainability, 11*(13), 3650.

Disclaimer/Publisher's Note