FROM CLAUSES TO CARBON: A SMART SUSTAINABLE PROCUREMENT TOOLKIT USING NLP AND LCA FOR INFRASTRUCTURE PROJECTS

Ariana Darvish[1] and Mehran Sepehri[2]

[1] *School of Built Environment, University of Reading, Reading, RG6 6UR, UK*

[2] *Construction Project Management, University of Portsmouth, Stanmore, London, HA7 3SB, UK*

Public procurement plays a pivotal role in achieving net-zero targets, yet current tendering practices often lack the tools to evaluate and operationalise sustainability commitments at scale. This paper presents a deployable Smart Sustainable Procurement Toolkit that enables procurement professionals to assess and benchmark sustainability clauses in infrastructure procurement documents. Developed and tested using UK railway sector data, the toolkit integrates Natural Language Processing (NLP) and Life Cycle Assessment (LCA) to semantically benchmark clauses against standards such as BREEAM, PAS 2080, and ISO 20400, and to quantify carbon impacts using verified Environmental Product Declarations (EPDs). Applied to over 300 procurement documents, the toolkit reveals patterns in credit prioritisation, actor-specific performance, and the weakening of sustainability across contractor tiers. While clients tend to issue strong, aligned clauses, suppliers and subcontractors often rely on vague or non-quantified language. By clause-level scoring, traceable carbon metrics, and an interactive dashboard for procurement, the toolkit enables transparent, enforceable, and data-driven procurement decisions. This work advances the use of NLP and LCA in construction management with a new model for semantic and carbon governance in sustainable procurement.

Keywords: BREEAM; environmental product declarations; life cycle assessment; natural language processing; sustainable procurement; semantic similarity

INTRODUCTION

Public procurement in the UK constitutes nearly a third of total public expenditure, with approximately £300bn spent annually on goods and services (Cabinet Office, 2022). This strategic purchasing power positions public infrastructure projects as a critical mechanism for delivering national sustainability and net-zero targets. Despite a growing policy momentum—from the Construction Playbook to the Transforming Infrastructure Performance: Roadmap to 2030—there remains a significant implementation gap in translating these frameworks into enforceable and measurable sustainability outcomes at the procurement stage (BEIS, 2021; Davies *et al.,* 2023).

Procurement for complex infrastructure like rail, is heavily document-driven. Early-stage procurement documents—including pre-qualification questionnaires (PQQs), specifications, and invitation to tender (ITT) materials—contain many of the embedded decisions that influence sustainability performance. Yet these texts are

[1] a.darvish@reading.ac.uk

Darvish, A and Sepehri, M (2025) From Clauses to Carbon: A Smart Sustainable Procurement Toolkit Using NLP and LCA for Infrastructure Projects, *In*: Thomson, C and Neilson, C J (Eds) Proceedings of the 41st Annual ARCOM Conference, 1-3 September 2025, Abertay University, Dundee, UK, Association of Researchers in Construction Management, 802-811.

typically unstructured, fragmented, and reliant on discretionary language, which undermines enforceability, comparability, and accountability (Olanrewaju *et al.,* 2024). The absence of standardised semantic classification for procurement clauses further impedes consistency in sustainability evaluation across actors and projects.

Recent studies highlight the potential of digital tools, particularly NLP, to address some of these challenges. Ding *et al. (*2022) demonstrate how NLP can automate semantic interpretation in construction contexts, while Gbadamosi *et al. (*2021) show its value in improving procurement decision-making through clause-level analysis. However, most applications remain conceptual, with limited deployment in real-time procurement environments or integration with enforceable sustainability benchmarks. More critically, the use of NLP to assess clause enforceability—e.g., mandatory versus suggestive language—remains underexplored in construction literature.

Simultaneously, LCA provides a robust, standardised method for quantifying the environmental impacts of materials and products. Despite this, LCA remains largely disconnected from early procurement processes due to barriers in data availability, integration, and tool accessibility (Waldman *et al.,* 2020). While studies such as Gachkar *et al. (*2025) explore the use of text-based algorithms to automate life cycle inventory analysis, real-time integration of LCA data—especially via EPDs—into procurement document analysis is still rare.

The literature also lacks consensus on how procurement clauses should be semantically classified and evaluated against established sustainability frameworks. For example, while frameworks like BREEAM and ISO 20400 offer high-level guidance, there is little empirical work on how their requirements are embedded (or diluted) across different procurement actors (Loosemore, 2016; McNutt and Casey, 2022). The lack of clause-level traceability across clients, contractors, and suppliers' limits accountability in delivering sustainable outcomes.

In response to these gaps, this research develops a deployable Smart Sustainable Procurement Toolkit that semantically benchmarks procurement clauses against sustainability standards and quantifies their carbon implications using verified LCA data. By aligning textual procurement data with environmental performance metrics, the toolkit addresses the dual challenge of unstructured clause language and the absence of standardised sustainability evaluation in procurement workflows.

METHOD

This study adopts a mixed-methods research design, integrating NLP and LCA into a clause-level benchmarking framework to evaluate the sustainability performance of procurement documents in infrastructure projects. The objective is to determine how well clauses in pre-construction procurement documents align with established sustainability standards and whether material-related language can be linked to verified carbon impact data.

The approach was implemented through the development of the Smart Sustainable Procurement Toolkit, which performs two core functions:

- Semantic benchmarking of clauses using NLP,
- Quantification of environmental impact using verified data from EPDs.

Data Sources

Two categories of data (Figure 1 and 2) were used in the analysis:

- Historical data: including BREEAM New Construction manuals, PAS 2080, ISO 20400, the Procurement Act 2023, NEC4 contracts, and the BRE GreenBookLive EPD database.
- Live procurement data: consisting of real Invitations to Tender (ITTs), Pre-Qualification Questionnaires (PQQs), technical specifications, and supplier ESG submissions obtained from a UK railway redevelopment via the Find a Tender portal, RISQS, and TfL Procurement Portal.

Documents were segmented into clauses and categorised by procurement actor (client, Tier 1 contractor, Tier 2 contractor, or supplier).

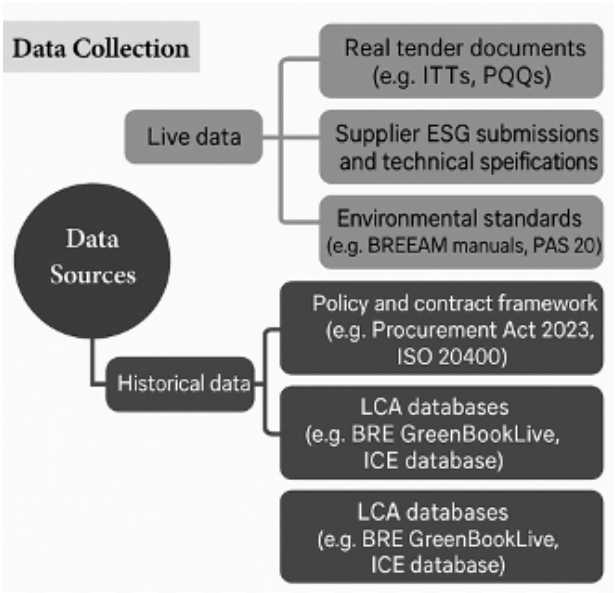

Figure 1: Phase 1 of Research Process: Data Collection

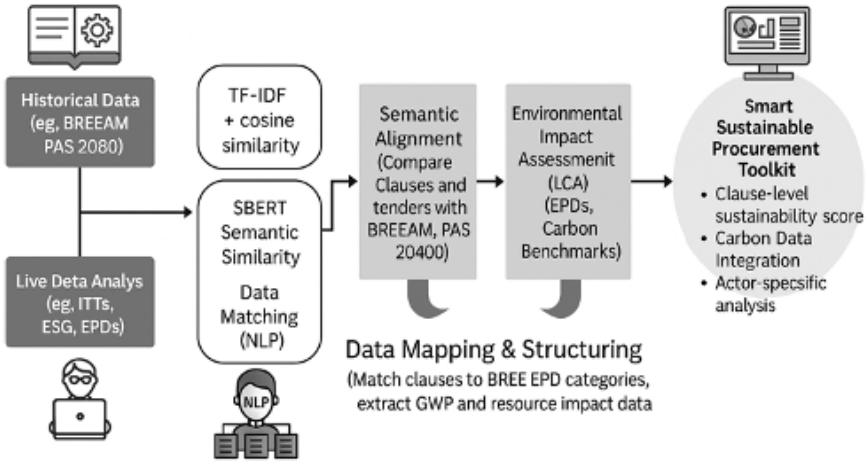

Figure 2: Phase 2 of Research Process: Analysis and Findings

NLP Tools and Justification

To assess semantic alignment between sustainability clauses and benchmark credit definitions (e.g., BREEAM Mat 01, ISO 20400), we employed two complementary NLP techniques:

- TF-IDF (Term Frequency-Inverse Document Frequency) was used to identify and weigh significant terms across documents, capturing how important a word is within a clause relative to the entire corpus.
- SBERT (Sentence-BERT) was applied to generate dense vector representations of clauses and benchmark definitions. SBERT captures contextual and semantic meaning at the sentence level, making it suitable for comparing full-length procurement clauses with structured sustainability credits (Reimers and Gurevych, 2019).

This dual method allows for both statistical term weighting (TF-IDF) and deep contextual understanding (SBERT), improving the reliability of clause-credit matching.

Clause Vectorisation and Similarity Scoring

To quantify semantic similarity, each clause and benchmark credit was vectorised using SBERT, and compared using cosine similarity:

$$\cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{||\vec{A}|| \cdot ||\vec{B}||}$$

where A is the clause vector and B is the benchmark credit vector. Cosine similarity measures the angular distance between two vectors, ranging from 0 (no similarity) to 1 (identical match).

Thresholds for alignment were defined based on empirical testing and aligned with prior studies (e.g., Ding *et al.,* 2022; Castelblanco *et al.,* 2021):

- Example of bullet points
- Strong alignment: cosine score ≥ 0.80
- Partial alignment: $0.50 \leq$ score < 0.80
- Weak alignment: score < 0.50

These thresholds were validated on a 50-clause pilot set, with manual review confirming the reliability of the 0.80 threshold for high-quality semantic matches.

Clause Modality and Enforceability

Each clause was analysed for modality—i.e., the strength of obligation—using rule-based identification of modal verbs. Clauses were classified as:

- Mandatory: use of "shall," "must," or "is required to"
- Suggestive: use of "should," "may," or "is encouraged to"

Modality was used to assess the enforceability of sustainability language across procurement tiers and linked to semantic and carbon performance.

Material Reference Detection and Carbon Quantification

Clauses with strong alignment to BREEAM Mat 01 were further screened for references to materials using Named Entity Recognition (NER) and rule-based keyword extraction. Phrases such as "steel frame," "timber panels," or "recycled concrete" were matched to corresponding product categories in the BRE GreenBookLive database.

To ensure traceability and validity:

- Only materials with third-party verified, ISO 14025-compliant EPDs were included.
- Carbon intensity data ($kgCO_2e/kg$) was extracted from modules A1-A3 (product and manufacturing stage), consistent with the scope of early-stage procurement decisions.

Clause Scoring Framework

Each clause was assigned a composite sustainability score based on three weighted components:

- Semantic similarity score (from cosine analysis) - 50%
- Carbon impact score (based on EPD data) - 50%
- Modality classification (used as an adjustment factor to penalise weak enforcement)

This scoring system enabled direct comparison of clauses across projects and procurement actors. Clauses with strong alignment, low-carbon material references, and mandatory framing received the highest scores (e.g., 4.5-5.0), while vague, suggestive, or non-quantifiable clauses scored below 2.0.

ANALYSIS AND DISCUSSION

4. Clause Alignment with Sustainability Standards

A key objective of the study was to assess how well sustainability-related clauses in procurement documents align with established benchmarks such as BREEAM, PAS 2080, and ISO 20400. Using NLP techniques—TF-IDF for term relevance and SBERT for semantic similarity—312 clauses from live and historical procurement documents (2020-2024) were analysed.

Clauses were categorised based on cosine similarity with BREEAM credit descriptors:

- Strong alignment (≥ 0.80)
- Partial alignment (0.50-0.79)
- Weak alignment (< 0.50)

Five key BREEAM credits were applied: Mat 01 (Life Cycle Impacts), Man 01 (Project Strategy), Mat 03 (Responsible Sourcing), Ene 01 (Operational Energy), and Pol 01 (Emissions). Among these, Mat 01 appeared most frequently, especially in client-issued documents, reflecting a dominant emphasis on early-stage material considerations. Man 01 and Mat 03 followed with strategic integration in procurement strategy and material sourcing. In contrast, Wat 01 (Water Efficiency) and Pol 01 (Refrigerants) were underrepresented, in contractor and supplier documents.

This uneven distribution highlights an important insight: sustainability credits closely linked to early procurement decision-making are more likely to be embedded, while those requiring long-term tracking or operational controls are deprioritised. This is consistent with prior findings from Olanrewaju *et al. (*2024), who argue that procurement documents act as "vehicles of sustainability intent"—where the presence, clarity, and structure of clauses directly influence project delivery outcomes.

However, intent alone does not guarantee delivery. Without strong semantic alignment and standardised framing, even well-meant clauses may lack enforceability or traceability downstream. This underlines the need for clause libraries and benchmarking tools, like the one developed in this study, to systematically assess how procurement texts reflect sustainability frameworks.

From a theoretical standpoint, this also reinforces Castelblanco et al.'s (2021) view of procurement as a semantic interface—not merely a legal or technical mechanism, but a medium through which environmental standards are operationalised across fragmented supply chains.

5. Modality and Enforceability of Clauses

Beyond semantic alignment, the enforceability of sustainability clauses—i.e., how strongly they are expressed—plays a critical role in translating procurement intent into action. To evaluate this, each clause was analysed for modality using rule-based detection of modal verbs. Clauses were classified as:

- Mandatory (e.g., shall, must)
- Suggestive (e.g., should, may)

Findings reveal a sharp decline in enforceability down the procurement hierarchy. 78% of client-issued clauses were mandatory, compared to 41% for Tier 1 contractors, and just 18% at the supplier level. This suggests a progressive weakening of sustainability commitments as documents cascade downstream.

This trend is further shaped by the type of BREEAM credit. Credits like Mat 01 (Life Cycle Impacts) and Man 01 (Project Strategy) were more frequently framed as enforceable requirements at the client level, whereas credits such as Pol 01 (Emissions) or Wat 01 (Water Efficiency) often appeared in suggestive language, especially among Tier 2 and supplier documents.

These findings align with McNutt and Casey (2022) and Loosemore (2016), who highlight that sustainability language is often diluted through subcontracting layers, contributing to a breakdown in accountability. From a theoretical perspective, this points to a critical weakness in procurement as a governance mechanism: while procurement documents may encode sustainability goals, they lack systemic enforcement structures to maintain obligation strength across all actor tiers.

ISO 20400, which promotes sustainable procurement practices, emphasizes the importance of clear, enforceable language to ensure consistent application. However, the observed variability suggests that compliance with such standards remains uneven. Without consistent modality across documents, even semantically aligned clauses can fail to drive measurable outcomes—particularly in relation to carbon-intensive materials or scope 3 emissions.

The Smart Sustainable Procurement Toolkit addresses this gap by automatically classifying clause modality alongside semantic benchmarking and carbon data. This feature allows procurement teams to detect where obligations are weakening, and to intervene early—either by strengthening downstream contract language or revising clause libraries for future tenders.

In sum, the enforceability of sustainability clauses is not only a technical issue but a theoretical one, exposing the limits of procurement documents as instruments of transformation unless supported by consistent linguistic and structural alignment.

6. Carbon Quantification Using EPDs

To move beyond qualitative alignment, the study introduces a clause-level carbon quantification process using verified data from EPDs. This component of the Smart Sustainable Procurement Toolkit links material-referencing clauses to ISO 14025-compliant product entries in the BRE GreenBookLive database, enabling a practical assessment of embedded carbon intent within procurement language.

Clauses aligned with BREEAM Mat 01 (Life Cycle Impacts) were screened using Named Entity Recognition (NER) and custom rule-based extraction to detect references to construction materials (e.g., "steel frame," "recycled concrete," "timber cladding"). A total of 102 Mat 01-aligned clauses were analysed:

- 46 clauses (45%) contained material references that could be matched to EPD categories.
- Carbon intensity values ($kgCO_2e/kg$) were retrieved from A1-A3 life cycle modules—the manufacturing stage most relevant to procurement-stage decision-making.

Table X presents examples of matched clauses, their semantic similarity scores, EPD carbon data, and final sustainability scores. For instance, a mandatory client clause referencing "recycled steel frame" received a composite score of 4.7 due to its strong semantic alignment (0.88) and low carbon intensity (0.92 $kgCO_2e/kg$). In contrast, a vague, suggestive supplier clause mentioning "durable concrete materials" without EPD linkage scored below 2.0.

This quantification process introduces a novel contribution to sustainable procurement practice: the ability to trace not just whether sustainability is mentioned, but to what extent it is measurable and enforceable through quantifiable product data. It supports prior calls by Waldman *et al. (*2020) and Gachkar *et al. (*2025) for more robust, automated integration of LCA data into early-stage construction decisions.

Theoretically, this work positions clause-based carbon scoring as a missing link between sustainability frameworks and procurement execution. Standards like PAS 2080 and ISO 20400 call for value-based, transparent procurement, but provide no operational guidance for evaluating unstructured text against environmental metrics. By filling this gap, the Toolkit enhances procurement's role as a carbon governance tool—not just through policy alignment, but through traceable, data-backed benchmarks embedded directly in contract language.

Furthermore, this approach reveals where semantic alignment fails to translate into carbon-conscious decision-making. Several clauses showed high similarity to BREEAM credits but lacked specific material references or EPD-compatible language, limiting their practical enforceability.

In sum, the integration of EPD-based quantification enables procurement teams to move from qualitative intent to quantitative verification, directly linking procurement language with measurable climate outcomes. This enhances both the accountability and granularity of sustainable procurement evaluations and suggests a new research avenue in clause-level LCA analytics for infrastructure governance.

7. Actor-Level Comparisons and Toolkit Use Case

A central function of the Smart Sustainable Procurement Toolkit is its ability to compare sustainability performance across procurement actors—clients, contractors (Tier 1 and Tier 2), and suppliers—based on clause alignment, enforceability, and carbon intent. These actor-level comparisons are visualised through the toolkit's interactive dashboard, which aggregates semantic scores, carbon intensities, and clause modalities for benchmarking.

Actor Comparison Insights

We studied five key BREEAM credits across actor tiers. The results indicate:

- Clients and delivery partners demonstrate the strongest semantic alignment, particularly for Mat 01 (Life Cycle Impacts), Man 01 (Project Strategy), and Ene 01 (Operational Energy).
- Tier 1 contractors exhibit moderate alignment, with noticeable gaps in enforceability and specificity, especially around Mat 03 (Responsible Sourcing).
- Suppliers and Tier 2 contractors show the lowest alignment across nearly all credits, and their clauses often lack carbon data or measurable commitments.

These findings underscore a declining gradient of sustainability ambition and precision as one moves downstream in the supply chain. While this pattern reflects the structural reality of construction contracting, it exposes a vulnerability in sustainability delivery: even if client-level clauses are robust, their downstream translation often weakens without standardised language, tools, or incentives.

This breakdown supports prior literature on fragmented accountability in procurement networks (Loosemore, 2016; Ghosh *et al.,* 2023), and extends Olanrewaju *et al. (*2024)'s argument that procurement systems must be audited not just at contract award but across the full document flow and actor hierarchy.

Toolkit Use Case and Practical Application

To address this fragmentation, the toolkit provides a visual and interactive interface where procurement teams can:

- View clause-level scores by actor and project phase.
- Identify underperforming credits or actor groups.
- Detect clauses lacking enforceability or carbon benchmarks.
- Compare clause wording side-by-side against high-scoring exemplars.

For example, a procurement officer preparing an Invitation to Tender (ITT) for a new rail station can use the dashboard (Figure 3) to:

- Review whether Mat 01 is already strongly embedded at the Tier 1 level.
- Flag supplier clauses that are suggestive rather than mandatory.
- Recommend the insertion of a pre-validated clause referencing low-carbon materials with linked EPDs.

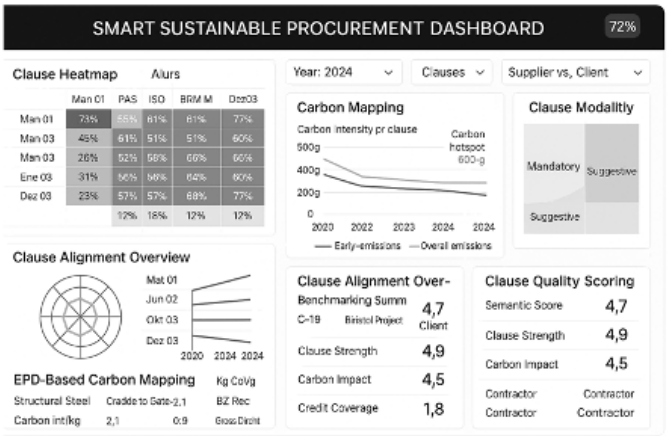

Figure 3: Smart Sustainable Procurement Dashboard: Clause-Level Insights and Carbon Benchmarking

This functionality transforms the toolkit from a research prototype into a practitioner-facing decision-support tool. It facilitates learning across projects, enforces

consistency in sustainability language, and supports alignment with frameworks like the Construction Playbook and PAS 2080. It enables procurement actors to not only comply with sustainability frameworks, but to do so with traceable, standardised, and verifiable documentation—a key missing mechanism in current infrastructure delivery.

CONCLUSIONS

This study presents a Smart Sustainable Procurement Toolkit that integrates NLP and LCA to benchmark sustainability clauses in infrastructure procurement documents. Developed using UK railway data, the toolkit aligns clauses with standards such as BREEAM, PAS 2080, and ISO 20400, assesses enforceability, and quantifies carbon impacts using verified EPDs.

Findings show strong clause alignment and enforceability at the client level, but a consistent decline downstream among contractors and suppliers. The toolkit highlights these gaps through actor-specific benchmarking and offers a practical decision-support interface for procurement teams.

Theoretically, the research contributes a clause-to-carbon model that positions procurement documents as tools for semantic and carbon governance. Practically, it enables traceable, standardised, and data-driven sustainability evaluations. Future work will extend the toolkit's scope, enhance generalisability across sectors, and integrate broader lifecycle data and international standards.

REFERENCES

Ajayi, S O, Oyedele, L O, Akinade, O O, Bilal, M, Owolabi, H A, Alaka, H A and Kadiri, K O (2015) Waste-efficient materials procurement for construction projects: A structural equation modelling of critical success factors, *Proceedings of the Institution of Civil Engineers - Waste and Resource Management*, **168**(2), 49-64.

BEIS (2021) *Transforming Infrastructure Performance: Roadmap to 2030 Department for Business, Energy and Industrial Strategy*, Available from: https://www.gov.uk/government/publications/transforming-infrastructure-performance-roadmap-to-2030 [Accessed 25 July 2025].

BRE (2022) *Greenbooklive: Directory of Certified Epds*, Available from: https://www.greenbooklive.com

BSI (2020) *PAS 2080: Carbon Management in Infrastructure*, London: British Standards Institution.

Cabinet Office (2022) *The Construction Playbook: Government Guidance on Sourcing and Contracting Public Works Projects and Programmes*, Available from: https://www.gov.uk/government/publications/the-construction-playbook [Accessed 25 July 2025].

Castelblanco, G, Guevara, J, Mesa, H and Sanchez, A (2021) Semantic network analysis of literature on public-private partnerships, *Journal of Construction Engineering and Management*, **147**(5), 04021033.

Choi, S W and Lee, E B (2022) Contractor's risk analysis of engineering procurement and construction (EPC) contracts using ontological semantic model and Bi-long short-term memory (LSTM) technology, *Sustainability*, **14**(11), 6938.

Davies, R, McKinlay, J and Zhang, C (2023) Assessing value-based procurement in UK infrastructure projects: A review of public-sector implementation, *Construction Management and Economic*s, **41**(2), 85-104.

Ding, Y, Ma, J and Luo, X (2022) Applications of natural language processing in construction, *Automation in Construction*, **136**, 104169.

Finkbeiner, M, Schau, E.M, Lehmann, A and Traverso, M (2010) Towards life cycle sustainability assessment, *Sustainability*, **2**(10), 3309-3322.

Gachkar, S, Gachkar, D, Ghofrani, E, Martínez, A G and Bahón, C A (2025) Text-based algorithms for automating life cycle inventory analysis in building sector life cycle assessment studies, *Journal of Cleaner Production*, **486**, 144448.

Gbadamosi, A Q, Oyedele, L O and Akinade, O O (2021) Big data and semantic analysis for decision support in construction procurement, *Advanced Engineering Informatics*, **48**, p.101314.

Ghosh, S, Oyedele, L O, Ajayi, A, Akinade, O and Bilal, M (2023) Overcoming barriers to sustainability in construction supply chains: A machine learning approach, *Journal of Cleaner Production*, **382**, 135122.

HM Government (2022) *Transforming Infrastructure Performance: Roadmap to 2030*, London: Infrastructure and Projects Authority.

ISO (2006) *ISO 14040: Environmental Management - Life Cycle Assessment - Principles and Framework*, Geneva: International Organisation for Standardisation.

ISO (2017) *Iso 20400: Sustainable Procurement – Guidance*, Geneva: International Organisation for Standardisation.

Loosemore, M (2016) Social procurement in UK construction projects, *International Journal of Project Management*, **34**(2), 133-144.

McNutt, P and Casey, D (2022) The enforceability of sustainability clauses in construction contracts: A legal analysis, *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, **14**(3), 04522015.

Moncaster, A M, Birgisdottir, H, Malmqvist, T, Nygaard Rasmussen, F, Houlihan Wiberg, A and Soulti, E, 2018 Embodied carbon measurement, mitigation and management within Europe, drawing on a cross-case analysis of 60 building case studies, *In*: *Embodied Carbon in Buildings: Measurement, Management and Mitigation*, 443-462.

Olanrewaju, O I, Enegbuma, W.I and Donn, M, 2024 Operational, embodied and whole life cycle assessment credits in green building certification systems: Desktop analysis and natural language processing approach, *Building and Environment*, 111569.

Oyedele, L O, Ajayi, S O and Kadiri, K O (2014) Use of recycled products in UK construction industry: An empirical investigation into critical impediments and strategies for improvement, *Resources, Conservation and Recycling*, **93**, 23-31.

Reimers, N and Gurevych, I (2019) Sentence-BERT: Sentence embeddings using Siamese BERT-networks, *In*: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 3982-3992.

Similä, J O, Mwesiumo, D E, Langseth, M and Haddara, M (2024) Exploring the antecedents and implications of data-driven decision-making in public procurement, *International Journal of Procurement Management*, **21**(2), 232-261.

UK Green Building Council (2022) Delivering Net Zero through Procurement, Available from: https://www.ukgbc.org

Waldman, B, Huang, M and Simonen, K (2020) Embodied carbon in construction materials: A framework for quantifying data quality in EPDs, *Buildings and Cities*, **1**(1), 1-18.