



AI-Based Construction Waste Reuse Matching Across Multiple Construction Projects.

Prof. Sonam Pattanaik¹, Mr. Soham Kale², Mr. Dhiraj Kakade³, Mr. Aditya Sonawane⁴

¹ Department of Artificial Intelligence and Machine Learning, ISBM College of Engineering, Nande, Pune, Maharashtra, India

^{2,3,4} Students, Department of Artificial Intelligence and Machine Learning, ISBM College of Engineering, Nande, Pune, Maharashtra, India

ABSTRACT :

Construction and demolition activities generate material streams that may retain value beyond the project in which they arise. At the same time, another project may need materials of a similar type, quantity, quality and delivery schedule. Identifying these opportunities manually is difficult because material compatibility, condition, location, transport distance, availability and project requirements must be considered together. This paper presents a proposed AI-based framework for matching reusable construction materials available at one project with demand at another. The framework standardizes project and material records, applies eligibility checks, represents relevant material and logistical attributes as comparable features, calculates a compatibility score, and ranks candidate supply-demand pairs. It is designed to provide an explanation for each recommendation, such as material-type agreement, quantity sufficiency, condition suitability, travel distance and timing. The design draws on research into digital workflows for construction-material reuse, machine learning for construction and demolition waste management, and material-passport data requirements. The contribution of this paper is a structured, explainable matching workflow that connects project-level supply and demand rather than predicting waste generation alone. This is a design proposal: no deployed system, real-world dataset, model training, or quantitative evaluation is claimed. A prototype evaluation plan is specified to test ranking relevance, constraint satisfaction, usability and potential environmental benefits using synthetic data initially and verified project data when available. The framework could support circular construction decisions, but any recommendation must remain subject to material inspection, engineering requirements and applicable safety standards.

Keywords: construction and demolition waste; material reuse; circular economy; machine learning; recommendation systems; compatibility scoring; geospatial analysis; explainable AI

INTRODUCTION :

Construction projects generate surplus and end-of-life materials, including bricks, concrete, steel, timber and glass. Some materials may be suitable for direct reuse or reuse after inspection and limited processing. However, the existence of surplus at one site does not automatically make it useful elsewhere: another project must need the same material, the material must meet its technical requirements, and the transfer must be feasible in terms of quantity, location and timing.

A central coordination problem is therefore to connect material supply with project demand. In a manual process, staff may need to compare separate inventories, specifications, site locations and schedules. A candidate match that appears suitable by material name alone may be rejected after checking dimensions, damage, contamination, certification or transport cost. These practical constraints mean that a useful digital tool must do more than search for similar labels.

This paper presents a proposed framework for matching material availability from one construction project to requirements at another. It combines structured data preparation, eligibility checks, feature-based compatibility scoring, ranking and human-readable explanations. The framework is intended as a decision-support tool: it does not certify materials or replace inspection by qualified personnel.

Problem Statement

There is no single, consistent comparison process in the proposed workflow for evaluating surplus material from multiple projects against the requirements of other projects while jointly considering material type, quantity, quality or condition, geographical location, transport distance, availability time and project demand. The problem addressed here is how to organize these attributes into a transparent workflow that can identify, score and rank feasible project-to-project reuse opportunities.

Objectives

The objectives of the proposed framework are to: (1) represent material availability and demand from multiple construction projects in a consistent format; (2) classify reusable materials and retain relevant condition and quality information; (3) compare supply and demand using material, quantity, location and time attributes; (4) exclude candidates that fail essential constraints; (5) calculate a compatibility score and rank feasible matches; (6) explain the main factors behind each recommendation; and (7) define an evaluation plan for testing ranking quality and estimating possible reuse benefits.

Scope and Contribution

The initial scope is a software prototype for common material categories such as bricks, concrete, steel, wood and glass, with a pilot geography limited to a city or region. Inputs may include project location, material type, quantity, condition, availability window, required quantity, required date and project-specific technical requirements. The proposed contribution is an explainable project-to-project matching workflow with explicit feasibility checks and a ranking stage. The paper does not claim a new trained algorithm, a completed production application, or experimentally demonstrated reductions in cost, waste or emissions.

LITERATURE SURVEY :

Digital systems for circular construction seek to make material information available at the point where reuse decisions are made. De Wolf et al. (2024) propose the D5 digital circular workflow—detection, disassembly, distribution, design and deployment—and report case-based use of digital tools to support reuse. Their work emphasizes that identifying materials is only one part of a larger process; distribution and coordination between supply and demand also matter. The proposed framework in this paper focuses on the narrower computational step of ranking candidate project-to-project matches from structured supply and demand records.

Gao et al. (2024) review machine-learning applications in construction and demolition waste management, analyzing research across multiple subfields and identifying continuing challenges. Their review is relevant because it shows that machine learning can contribute to waste-related prediction and management, while also drawing attention to dataset quality, deployment and generalization. Predicting how much waste a project will generate, however, is not equivalent to selecting the best recipient for a known surplus batch.

Cha et al. (2024) investigate machine-learning models for predicting demolition-waste generation using building and waste-characteristic information. For the matching problem considered here, a prediction model could eventually estimate availability or expected quantities, but it would need to be combined with a separate compatibility and logistics stage.

Çetin et al. (2023) examine data requirements and availability for material passports in existing buildings. Their study identifies gaps relating to material composition, hazardous content, condition assessment and reuse or recycling potential. Such findings matter directly to a matching system: inaccurate or missing material records can make a high numerical similarity score misleading. A recommendation should therefore expose data completeness and uncertainty rather than hide them.

Taken together, these studies suggest a useful distinction between (a) identifying and documenting materials, (b) predicting waste or material quantities, and (c) matching documented supply to actual project demand. The proposed framework addresses the third task while relying on the first as a prerequisite.

Research Gap and Design Rationale

The reviewed sources address digital reuse workflows, waste-management machine learning and material-passport data. The present seminar materials do not establish that an existing system is unable to perform project matching, and this paper therefore does not claim a universal absence of prior solutions. Instead, the design rationale is a practical integration need: to express material and project constraints in a consistent matching pipeline, rank candidate pairs, and make the reasons for each ranking visible to a user. The proposed work is a design contribution that requires prototype implementation and empirical validation before its effectiveness can be assessed.

PROPOSED METHODOLOGY :

The workflow is organized into seven stages: data intake; data validation and standardization; candidate generation; feasibility filtering; feature construction; compatibility scoring and ranking; and explanation with human review. Figure 1 summarizes the proposed flow.

Input Data and Feature Representation

Each project may provide supply records for materials available for reuse and demand records for materials required. A supply record should include a project identifier, material category, available quantity and unit, condition or grade, dimensions or specification where relevant, location, availability start and end dates, inspection status and supporting evidence. A demand record should include a receiving project identifier, required material category, quantity and unit, minimum acceptable condition or grade, required specification, destination location and required delivery window.

Before matching, units and category labels must be standardized. Quantities expressed in incompatible units should not be compared until a valid conversion is available. Missing values should be marked as unknown rather than silently interpreted as a match. Where a critical property such as structural grade or contamination status is unknown, the system should flag the record for review or exclude it from automatic recommendations, depending on the material and intended use.

Table 1. Core information for matching

Feature	Supply-side example	Demand-side example	Matching role
Material identity	Brick, steel, timber	Required category	Mandatory category compatibility
Quantity	500 units or 2 tonnes available	300 units or 1 tonne required	Check sufficiency and partial allocation
Condition / quality	Inspected, grade, damage notes	Minimum acceptable condition	Safety and suitability constraint
Specification	Size, grade, dimensions	Required dimensions or grade	Technical compatibility

Location	Supply-site coordinates or address	Receiving-site location	Estimate route distance or travel cost
Availability / timing	Available from date to date	Required by date	Check temporal feasibility
Evidence completeness	Inspection record, photos, provenance	Acceptance criteria	Confidence and human-review flag

Data Validation and Candidate Generation

The system first validates mandatory fields, units, date ranges and category labels. Records with missing or contradictory information should be returned for correction or marked as uncertain. It then generates candidate pairs by comparing supply records with demand records of compatible material categories. Candidate generation may use category filters and a geographical radius to avoid scoring every possible pair when the number of projects grows. The radius should be configurable because a long-distance transfer may be reasonable for high-value materials but uneconomical for low-value bulk materials.

Feasibility Filtering

A high score should not override an essential constraint. Before ranking, the system should apply feasibility rules, such as compatible material category, adequate or allocatable quantity, required technical specification, acceptable condition, valid availability dates and any user-defined maximum transport distance. These rules should be material-specific. For example, an item that requires structural certification should not be recommended for a structural use solely because its category and location match. If only part of the demand can be met, the interface should state that the match is partial and should prevent the same supply quantity from being allocated to multiple demands beyond the available amount.

Compatibility Scoring

For candidates that pass the feasibility stage, the system can compute a normalized score from several factors. One possible initial design is a weighted linear score:

$$C(s,d) = 100 \times [w_m M + w_q Q + w_c K + w_l L + w_t T + w_e E]$$

where $C(s,d)$ is the compatibility score for supply record s and demand record d ; M represents material-category and specification compatibility; Q represents quantity fit; K represents condition suitability; L represents location or transport suitability; T represents timing compatibility; and E represents evidence completeness. Each component is scaled to the interval $[0,1]$, and the weights are non-negative with their sum equal to 1. Mandatory conditions are handled by feasibility filters rather than compensated for by other high scores.

The score is a proposed design choice, not a validated formula. Weights should be selected with domain-expert input and tested through sensitivity analysis. A future version could compare a transparent weighted baseline with learning-to-rank or other supervised methods if a sufficiently large, labelled dataset of accepted and rejected matches becomes available. Without such data, describing the system as a trained predictive model would be inaccurate.

Ranking and Recommendation Explanation

After scoring, feasible candidates are sorted from highest to lowest compatibility. Each result should display the score alongside the factors that influenced it. An explanation might state that the material category and specification match, the quantity covers the requested amount, the delivery window overlaps, and the estimated route is within the selected threshold. It should also identify limitations such as incomplete inspection evidence, partial quantity coverage or a distance estimate based on approximate coordinates. A numerical score should be treated as a prioritization aid, not a guarantee that the material is safe or suitable.

Figure 1. Proposed workflow

Project supply and demand records → validation and standardization → candidate generation → mandatory feasibility checks → feature construction → compatibility scoring → ranked, explainable recommendations → human inspection and approval

Proposed System Architecture

A prototype can be organized into five logical components. (1) A data-entry layer captures project and material records. (2) A storage layer maintains project, supply, demand and inspection information in a relational database. (3) A processing layer standardizes records and performs validation and candidate filtering. (4) A matching engine computes compatibility scores, ranks candidates and creates explanation fields. (5) A user interface presents the ranked list, filters and evidence needed for review. A basic implementation could use Python, Pandas, scikit-learn, GeoPy or GeoPandas, SQLite or MySQL, and Streamlit, consistent with the technologies listed in the supplied synopsis. Deep-learning tools are optional and are not required for the initial rule-based or weighted-score baseline.

Technology Requirements

The supplied synopsis lists Windows 10/11, Python 3.10 or later, VS Code or Jupyter Notebook, NumPy, Pandas, scikit-learn, optional PyTorch or TensorFlow, GeoPy or GeoPandas, NetworkX, Matplotlib or Seaborn, Streamlit, SQLite or MySQL, and Git/GitHub. These are proposed tools rather

than evidence that a working implementation has been completed. The minimum viable prototype can be built with Python, Pandas, a database, a geospatial distance utility and a simple interface.

PROTOTYPE EVALUATION PLAN :

Because the supplied project documents do not report a completed implementation or measured results, evaluation is proposed rather than reported as completed. The first phase should use a synthetic dataset with known supply–demand relationships to test the pipeline, including edge cases such as insufficient quantity, incompatible specifications, missing inspection information, overlapping dates and duplicate allocation. A later phase should use appropriately authorized, de-identified project data and domain-expert labels.

Evaluation should include: (1) constraint satisfaction—the proportion of displayed recommendations that pass mandatory rules; (2) ranking relevance—precision at k or normalized discounted cumulative gain if expert-labelled relevance judgements are available; (3) quantity allocation integrity—whether total allocated supply exceeds available quantity; (4) distance and timing validity—whether route and availability constraints are applied consistently; (5) explanation usefulness—reviewer assessment of whether the displayed reasons are understandable and faithful to the score; and (6) operational measures such as response time as the number of projects grows. Environmental and financial benefits should be estimated only when the system has defensible records of actual reuse, avoided procurement, transport and disposal assumptions.

A baseline comparison could compare the proposed ranking with manual review or a simple nearest-site strategy. No numerical performance result should be stated until the tests have been conducted and documented.

Table 2. Proposed evaluation plan

Evaluation dimension	Method	Evidence needed
Constraint satisfaction	Audit every recommended pair against mandatory rules	Test cases and rule-check logs
Ranking relevance	Precision@ k or NDCG against expert relevance labels	Independent expert judgements
Quantity integrity	Reconcile all allocations against supply totals	Allocation ledger
Temporal and geographic feasibility	Check dates and route-distance calculations	Valid dates and geocoded locations
Explainability	Ask reviewers to assess clarity and correctness of explanations	Structured reviewer feedback
Scalability	Measure processing time at increasing record counts	Repeatable benchmark setup
Environmental / cost estimates	Compare verified reuse outcomes with a stated baseline	Actual transfers and documented assumptions

DISCUSSION :

The framework is intended to support the coordination step between available material and project demand. Its main design strengths are the separation of hard feasibility constraints from ranking, the explicit treatment of data quality, and the display of reasons behind each recommendation.

The approach also has limitations. First, the score depends on the quality and completeness of material records. Second, condition cannot always be inferred reliably from text fields or photographs; physical inspection and appropriate testing may be necessary. Third, distance is only a proxy for transport cost and environmental impact: vehicle type, load, road conditions, handling, storage and return trips can change the result. Fourth, projects may use different naming conventions, units and documentation practices. Fifth, commercial sensitivity may limit the sharing of inventory, cost and schedule information. These issues should be addressed through standard data fields, access controls, explicit uncertainty indicators and a process for human approval.

The framework should not be interpreted as an autonomous material-certification system. It can identify promising candidates for review, but it cannot by itself establish compliance with building codes, structural design requirements or material-specific safety rules. Any future deployment should define who verifies material condition, who accepts responsibility for the transfer and what documentation accompanies the material.

Operationally, the system should preserve the source record for each recommended material batch. Users should be able to inspect the original description, availability date, supporting evidence, and edits made during validation. A recommendation should distinguish a confirmed match from a tentative candidate, particularly when information is incomplete or the receiving project has not approved a substitute. A traceable decision history can help teams review why a transfer was considered, rejected, delayed or completed. Such records may support later evaluation when collected with permission and protected against unauthorized

Potential Applications

Potential use cases include connecting surplus bricks or timber from one project with another project that can use them; identifying steel or other materials that may be reused after appropriate inspection; helping project managers compare nearby sources before procuring new material; and maintaining a record of why a match was accepted or rejected.

Limitations and Ethical Considerations

The prototype would depend on accurate and timely information from participating projects. Location and commercial data should be collected only for a defined purpose, access should be role-based, and data-sharing agreements should specify permitted uses and retention. The system should avoid presenting incomplete data as certainty, and it should record the assumptions used for distance, quantity and benefit estimates. Human review remains necessary where safety, legal compliance or engineering judgement is involved.

CONCLUSION AND FUTURE WORK :

This paper has described a proposed explainable framework for matching construction-material supply from one project with demand at another. The workflow combines standardized records, candidate generation, feasibility filtering, feature-based compatibility scoring, ranking and recommendation explanations. The design addresses a practical coordination problem that is related to, but distinct from, waste-generation prediction and material-passport data management. Its intended value lies in organizing multiple constraints into a transparent decision-support process.

The framework has not yet been validated through a deployed system or quantitative experiments, so reductions in waste, cost, transport or emissions remain potential benefits rather than demonstrated results. Future work should implement a prototype, create a carefully documented test dataset, obtain expert relevance labels, compare the ranking against transparent baselines, and test the system with authorized project data. Further extensions may include integration with building information modelling, digital material passports, live inventory systems, mobile site access, transport-cost optimization and sustainability reporting.

REFERENCES :

1. Cha, G.-W., Park, C.-W., & Kim, Y.-C. (2024). Optimal machine learning model to predict demolition waste generation for a circular economy. *Sustainability*, 16(16), 7064. <https://doi.org/10.3390/su16167064>
2. Çetin, S., Raghu, D., Honic, M., Straub, A., & Gruis, V. (2023). Data requirements and availabilities for material passports: A digitally enabled framework for improving the circularity of existing buildings. *Sustainable Production and Consumption*, 40, 422–437. <https://doi.org/10.1016/j.spc.2023.07.011>
3. De Wolf, C., Byers, B. S., Raghu, D., Gordon, M., Schwarzkopf, V., & Triantafyllidis, E. (2024). D5 digital circular workflow: Five digital steps towards matchmaking for material reuse in construction. *npj Materials Sustainability*, 2, Article 36. <https://doi.org/10.1038/s44296-024-00034-8>
4. Gao, Y., Wang, J., & Xu, X. (2024). Machine learning in construction and demolition waste management: Progress, challenges, and future directions. *Automation in Construction*, 166, 105380. <https://doi.org/10.1016/j.autcon.2024.105380>