Project 2524 | September 2026

SJSU SAN JOSÉ STATE UNIVERSITY

MTI MINETA TRANSPORTATION INSTITUTE

AI-Based Environmental Code Checking Tool for Sustainability Best Management Practices of Infrastructure Construction Projects

Joseph J. Kim, PhD
Pooja D. Chavan, MS

California State University Transportation Consortium

CALIFORNIA STATE UNIVERSITY LONG BEACH

MINETA TRANSPORTATION INSTITUTE

Founded in 1991, the Mineta Transportation Institute (MTI), an organized research and training unit in partnership with the Lucas College and Graduate School of Business at San José State University (SJSU), increases mobility for all by improving the safety, efficiency, accessibility, and convenience of our nation's transportation system. Through research, education, workforce development, and technology transfer, we help create a connected world. MTI leads the California State University Transportation Consortium (CSUTC) funded by the State of California through Senate Bill 1 and the Rail Resilience to Severe Weather: Training and Research Program (RR2SW) funded by the Federal Railroad Administration. MTI focuses on three primary responsibilities:

Research

MTI conducts multi-disciplinary research focused on surface transportation that contributes to effective decision making. Research areas include: active transportation; planning and policy; security and counterterrorism; sustainable transportation and land use; transit and passenger rail; transportation engineering; transportation finance; transportation technology; and workforce and labor. MTI research publications undergo expert peer review to ensure the quality of the research.

Education and Workforce Development

To ensure the efficient movement of people and goods, we must prepare the next generation of skilled transportation professionals who can lead a thriving, forward-thinking transportation industry for a more connected world. To help achieve this, MTI sponsors a suite of workforce development and education opportunities. The Institute supports educational programs offered by the Lucas Graduate School of Business: a Master of Science in Transportation Management, plus graduate certificates that include High-Speed and Intercity Rail Management and Transportation Security Management. These flexible programs offer live online classes so that working transportation professionals can pursue an advanced degree regardless of their location.

Information and Technology Transfer

MTI utilizes a diverse array of dissemination methods and media to ensure research results reach those responsible for managing change. These methods include publication, seminars, workshops, websites, social media, webinars, and other technology transfer mechanisms. Additionally, MTI promotes the availability of completed research to professional organizations and works to integrate the research findings into the graduate education program. MTI's extensive collection of transportation-related publications is integrated into San José State University's world-class Martin Luther King, Jr. Library.

Report 26-29

AI-Based Environmental Code Checking Tool for Sustainability Best Management Practices of Infrastructure Construction Projects

Joseph J. Kim, PhD
Pooja D. Chavan, MS

September 2026

A publication of the
Mineta Transportation Institute
Created by Congress in 1991

College of Business
San José State University
San José, CA 95192-0219

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
26-29		

4. Title and Subtitle	5. Report Date
AI-Based Environmental Code Checking Tool for Sustainability Best Management Practices of Infrastructure Construction Projects	September 2026
	6. Performing Organization Code
7. Authors Joseph J. Kim, PhD Pooja D. Chavan, MS	**8. Performing Organization Report** CA-MTI-2524
9. Performing Organization Name and Address Mineta Transportation Institute College of Business San José State University San José, CA 95192-0219	**10. Work Unit No.** **11. Contract or Grant No.** SB1-SJAUX_2023-26
12. Sponsoring Agency Name and Address State of California SB1 2017/2018 Trustees of the California State University Sponsored Programs Administration 401 Golden Shore, 5th Floor Long Beach, CA 90802	**13. Type of Report and Period Covered** **14. Sponsoring Agency Code**

15. Supplemental Notes

10.31979/mti.2026.2524

16. Abstract

Infrastructure construction and operation are major contributors to global greenhouse gas emissions, which account for approximately one-third of global CO_2 emissions and add to long-term environmental and air quality challenges. Sustainability rating systems such as Envision for infrastructure construction projects provide structured guidance to help teams reduce these impacts, but verifying compliance requires reviewing large volumes of project documents by hand, making the process time consuming and difficult to scale. Therefore, this research develops an AI-based environmental code checking tool that automates the interpretation of project documents and evaluates compliance on lifecycle GHG assessment under Credit CR1.2 of Envision—in other words, determining whether a project successfully reduces greenhouse gas emissions throughout its lifecycle. This research integrates natural language processing, life-cycle assessment, and Envision-based scoring into a unified analytical pipeline. Using a custom spaCy-based Named Entity Recognition (NER) model, the authors trained the system on a representative case study based on real-world empirical data of infrastructure to extract project metadata, material quantities, and activity data from PDF documents. Extracted values were standardized and processed through an LCA module to calculate total, annualized, and intensity-based emissions, which were then benchmarked against baseline scenarios to determine Envision performance levels. A Streamlit web application was created to enable rapid document review with automated extraction, emissions calculation, visualization, and credit scoring. Based on a bridge case study, the system achieved high extraction accuracy with an F1-score of 95.6%, completed assessments in under five seconds, and correctly classified the project as "Improved" with a 17.5% GHG reduction. These findings demonstrate that AI can **significantly reduce the time and effort required to evaluate sustainability performance while improving consistency and scalability**, supporting broader adoption of standardized digital documentation, early-stage LCA integration, and AI-assisted verification practices in infrastructure construction projects.

17. Key Words	18. Distribution Statement
Artificial intelligence, sustainable transportation, life cycle analysis, infrastructure, greenhouse gases	No restrictions. This document is available to the public through The National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	35	

Form DOT F 1700.7 (8-72)

DOI: 10.31979/mti.2026.2524

Mineta Transportation Institute
College of Business
San José State University
San José, CA 95192-0219

Email: mineta-institute@sjsu.edu

transweb.sjsu.edu/research/2524

Acknowledgments

The writers of this report thank Dr. Hamid Rahai at California State University, Long Beach, and Dr. Hilary Nixon, Deputy Executive Director of Mineta Transportation Institute at San José State University, for their kind guidance. The authors also thank Aissatou Diallo, student assistant, for helping to prepare the draft of the manuscript using the outcomes provided by the authors.

Contents

List of Figures

List of Tables

1. Introduction

1.1 Background and Motivation

The concept of sustainability in highways, as outlined by the Federal Highway Administration (FHWA), places equal importance on environmental, economic, and social values. This definition indicates that sustainable highways should prioritize safety, mobility, environmental protection, livability, asset management, and effective cost management throughout their life cycle (FHWA, 2018). With the urgent need to care for the world's limited resources, the construction industry has increasingly embraced sustainable development (Reeder, 2010). Transportation ranks as the second-highest energy-consuming sector in the United States, following electric power, industrial, residential, and commercial sectors (EIA, 2018). Taking swift action to reduce energy consumption in transportation is crucial for ensuring climate resilience and a sustainable future for generations to come. Ensuring the fair distribution of resources among all people acknowledge their right to equitable access to materials, land, energy, water, and environmental quality (Kim & McCarthy, 2022).

To address sustainability concerns within the transportation sector, the significance of best management practices (BMPs) has reached unprecedented levels. Therefore, it is urgent and critical to incorporate sustainability BMPs rooted in regulatory compliance throughout the entire life cycle of infrastructure construction projects. Despite the availability of various rating systems nationwide, the absence of thorough and comprehensive research in the realm of regulatory compliance poses a challenge. This gap hinders project teams from making well-informed decisions regarding the selection of sustainability BMPs for a specific project.

Recently, artificial intelligence (AI) technologies have been widely adopted as research methods in transportation and infrastructure construction research communities far more than before because the size and complexity of the projects increases exponentially and because the need to address the critical and contemporary sustainability issues is vital. Machine learning (ML) is a subset of AI that focuses on algorithm and model development in which computers can learn from data and improve their performance over time without being explicitly programmed (Chen & Zhang, 2014). AI technologies and ML models provide support for construction management decisions and are being applied in several environmental areas, including air pollution (Mo et al., 2019; Ashayeri et al., 2021; Lin et al., 2018; Bashir Shaban et al., 2016; Huang et al., 2021; Kleine Deters et al., 2017; Babu Saheer et al., 2022; Zhou et al., 2018; Suárez Sánchez et al., 2013; Yeganeh et al., 2012), waste management (Gue et al., 2022; Al-Ruzouq et al., 2022; Solano Meza et al., 2019), noise pollution (Renaud et al., 2023), infrastructure management (Sousa et al., 2022; Yu et al., 2021; Zou et al., 2023; Xu et al., 2022; Cassottana et al., 2022; Cha et al., 2023; Zhu et al., 2021; Ekici et al., 2021; Renaud et al., 2023), and life cycle assessment and rainfall prediction (Koyamparambath et al., 2022; Al Duhayyim et al., 2022), as well as construction management decisions (Gue et al., 2022; Pan & Zhang, 2021; Naumets & Lu, 2021; Pan et al., 2022). AI technologies are also being applied in several construction areas, including construction safety (Fang et al., 2018a, Fang et al., 2018b, Fang et al., 2018c, 2019; Wu & Cai, 2019), road construction surveys (Zhang & Yang, 2016; Wu et al., 2019), bridge construction inspections (Deng et al., 2020; Dorafshan & Azari, 2020; Zhang & Yang, 2020), and on-site construction operation monitoring (Fang et al., 2018d; Guo et al., 2020).

Gao et al. (2019) and Zhang and Shi (2020) reviewed machine learning for life cycle management, while Xu et al. (2021) reviewed deep learning to provide different research directions. These studies combined algorithms and application scenarios to solve their problems. Although AI technologies and ML models are becoming increasingly popular and have been successfully applied to these areas, their use has not been well applied in addressing environmental regulatory compliance for infrastructure construction projects. Therefore, there is a significant need and motivation to explore how AI can be incorporated into sustainability BMPs regarding site, water and wastewater, energy, material, and environmental problems and to further address the immediate challenges of meeting regulatory compliances for California's infrastructure construction projects.

1.2 Problem Statement

Construction and operation of infrastructure assets are major contributors to global greenhouse gas (GHG) emissions, resource depletion, and others have long-term environmental impacts. Recent analyses have shown that the construction industry is responsible for approximately 34% of global CO_2 emissions (UNEP, 2025). Similarly, the Climate Group (2025) reports that 40% of global GHG emissions originate from buildings, emphasizing the urgency of transitioning towards low-carbon infrastructure systems. To address these impacts, sustainability assessment frameworks, such as U.S. Green Building Council (USGBC)'s Leadership in Energy and Environmental Design (LEED), Building Research Establishment Environmental Assessment Method (BREEAM), and Envision, have been developed to guide stakeholders in adopting environmentally responsible design and construction practices.

Among these frameworks, Envision was created by the Institute for Sustainable Infrastructure (ISI, 2018) to provide a comprehensive methodology to evaluate sustainability across a range of infrastructure types, including bridges, highways, water systems, and energy facilities. The Envision rating standard sets the standard for what constitutes sustainable, resilient, and equitable infrastructure, and it provides a common language for collaboration and clear communication. Envision consists of five categories comprising 64 credits that collectively define sustainability in infrastructure systems: *Quality of Life*, *Leadership*, *Resource Allocation*, *Natural World*, and *Climate/Resilience* (ISI 2018). One of the credits, "CR1.2: Reduce Greenhouse Gas Emissions," specifically targets lifecycle GHG reductions relative to a baseline and directly supports climate mitigation objectives, making it increasingly relevant in the context of global carbon-neutrality commitments.

Despite the rigor of the Envision framework, its compliance verification process remains predominantly manual. Assessors are required to review extensive project documentation, extract material and operational data, and compute lifecycle performance metrics such as GHG. This reliance on manual procedures introduces several key limitations:

- Time-intensiveness: Reviewing and validating data across hundreds of pages demands substantial effort and slows the assessment process.

- Subjectivity: Manual interpretation increases the potential for inconsistencies between assessors, affecting the reliability of evaluations.

- Limited scalability: As project volume increases, applying Envision consistently across multiple infrastructure projects becomes impractical without automation.

A prior review study examined automation in environmental and regulatory compliance (Sečkár et al., 2025). The study found that early efforts focused on Life Cycle Assessment (LCA) tools, such as SimaPro, OpenLCA, and GaBi; however, these systems require structured, well-formatted inputs and, therefore, do not resolve the fundamental challenge of extracting relevant information from unstructured project documentation. More recent studies highlighted the use of Natural Language Processing (NLP) and Building Research Establishment Environmental Assessment Method ML in related domains (Al Qady & Kandil, 2010; Zhang & El-Gohary, 2022; Eken et al., 2023, Thimm, 2023).

For instance, Zhang and El-Gohary (2022) developed automated compliance checking that combines NLP, AI, and Building Information Modeling (BIM) to process regulatory requirements and detect non-compliance, underscoring the growing maturity of text analytics–based compliance systems. Eken et al. (2023) explored automated contract analysis through NLP-driven summarization and risk identification, demonstrating the viability of text extraction and interpretation for complex technical documents. Thimm (2023) proposed an NLP-based scoring method to evaluate the relevance of environmental regulatory announcements. However, despite these advances, the literature reveals limited work integrating AI-based document interpretation with the Envision sustainability framework, particularly for infrastructure-scale applications where documentation is lengthy, diverse, and technically complex.

To address these gaps, we propose and develop an AI-based environmental code compliance checking system that integrates Named Entity Recognition (NER) with LCA-based GHG calculations. First, we design and train a domain-specific NER model to extract project parameters from unstructured PDF reports. Second, we develop an automated pipeline to normalize values, compute embodied and operational emissions, and map results to Envision

CR1.2's scoring criteria, consistent with ISI guidance. Third, we implement a prototype Streamlit application supporting end-to-end verification. Finally, we evaluate the proposed system via a bridge case study demonstrating accuracy, efficiency gains, and scalability. The proposed approach can enhance existing studies by bridging AI-driven NLP with infrastructure sustainability assessment, moving toward automated, reproducible, and scalable compliance verification under the Envision framework.

The objective is to develop an AI-based compliance verification system capable of automating the evaluation of infrastructure projects under the Envision sustainability framework, with particular emphasis on Credit CR1.2 to reduce GHG emissions. Existing verification processes rely extensively on manual examination of project reports, where a workflow is labor-intensive, prone to interpretation inconsistencies, and susceptible to data-handling errors. To overcome the limitations of existing verification processes, the proposed system integrates an NLP workflow by featuring a custom-trained NER model and extracting critical project parameters, including material quantities, units, energy consumption, operational lifespan, and baseline emissions, directly from unstructured PDF documentation.

To demonstrate the viability of the proposed AI model, this report presents the framework and preliminary results focusing specifically on Envision Credit CR1.2: Reduce Greenhouse Gas Emissions. While CR1.2 is only one of 64 credits within the Envision rating system, targeting lifecycle GHG reductions serves as a strategic pilot to launch our tool's Life Cycle Assessment (LCA) module, which is the core computational engine of the application. Although integrating all 64 credits remains a complex, ongoing effort, this initial implementation demonstrates the significant promise of using natural language processing (NLP) to automate data extraction for environmental compliance. Ultimately, these preliminary results validate the core architecture and establish a robust baseline for future, full-scale implementation. The AI model was tested using a representative case study framework compiled from real-world empirical data, including project metadata, material inventories, and activity metrics, sourced directly from an authentic Bridge Lifecycle GHG Report.

2. Research Methodology

We adopted a design science research methodology to develop, implement, and evaluate an AI-based environmental compliance checking system aligned with the Envision sustainability framework. We aimed to automate lifecycle GHG assessment, reduce human error in data extraction and processing, and improve the transparency and reproducibility of credit evaluations. The methodology integrates case study analysis, NLP, LCA, and Envision-based scoring into a unified analytical pipeline. We used the Life-Cycle Assessment (LCA) method by Dequidt (2012) that enables quantification of total emissions and comparison against baseline scenarios.

To contextualize system design and construct training datasets, we used the Sixth Street Viaduct Seismic Improvement Project (ISI, 2022), an infrastructure project that achieved an Envision Platinum award from the ISI based on its performance in community benefits, safety enhancement, urban connectivity, and long-term resiliency. By examining publicly available ISI project documentation, we identified the data types required for Envision verification, including material quantities, energy use, projected operational life, and other sustainability parameters. These elements informed the construction of domain-specific training corpora for the NLP pipeline.

In addition to analyzing real-project documentation, a Bridge Lifecycle GHG Report was used to evaluate the system's extract–transform–compute fidelity under controlled conditions. This dataset incorporated typical infrastructure parameters, including concrete and steel quantities, fuel consumption, and baseline lifecycle GHG emissions, to enable systematic testing of the NLP extraction workflow, unit normalization procedures, and LCA computation modules without reliance on a single real-world source. These complementary data sources provide a rigorous and repeatable foundation for assessing the effectiveness of AI-enabled compliance automation within the Envision framework.

To simulate realistic reporting conditions, the Bridge Lifecycle GHG Report included representative project metadata (project name, location, operational lifespan, and baseline emissions), as well as material and activity inventories (concrete, steel, diesel, electricity, and transport-related metrics). The variability, formatting inconsistencies, and unstructured nature of such documentation underscore the need for automated data-extraction techniques. To address this challenge, a custom NER model was developed using the spaCy library. We used Python to execute our workflow and subsequent analyses. The model was trained in diverse textual patterns to identify three primary entity groups essential to lifecycle emissions analysis:

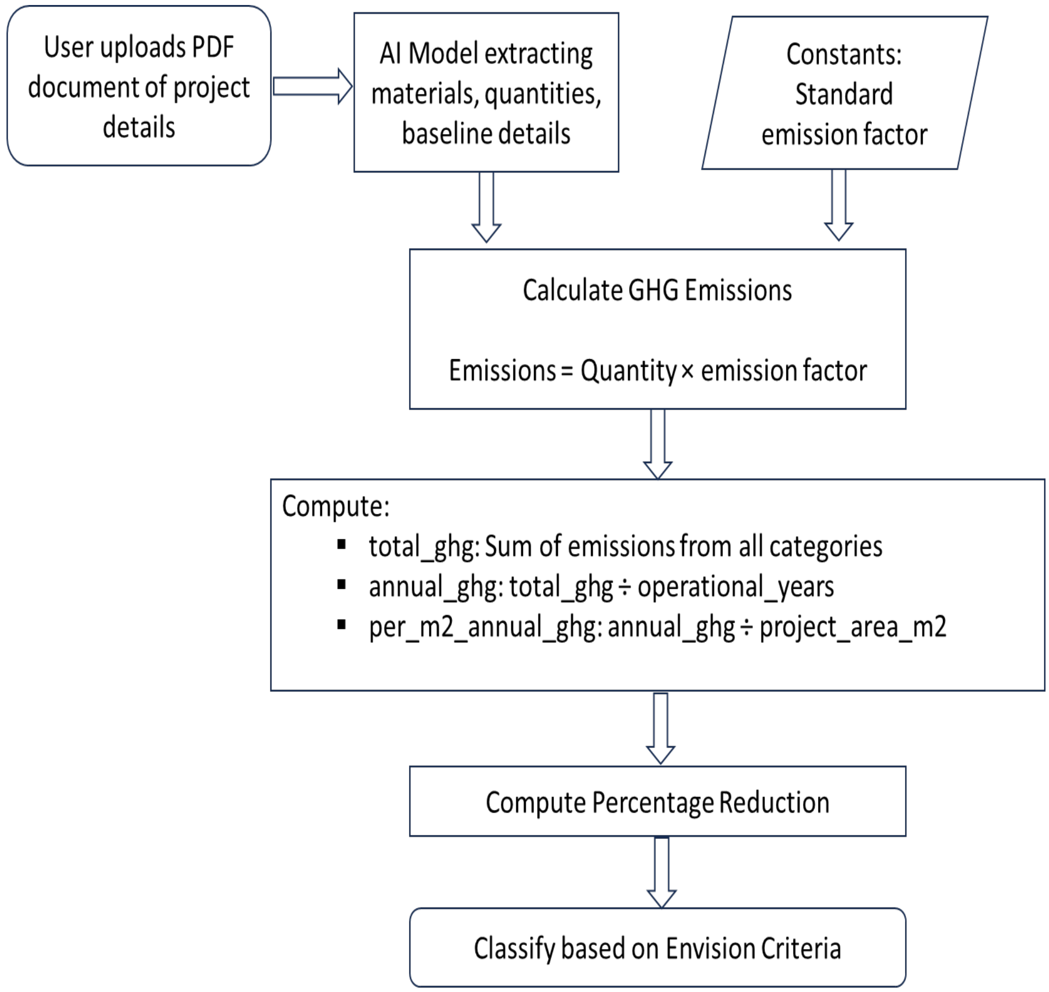
1. Project metadata: project name, location, operational life, baseline GHG values
2. Materials and activities: concrete, steel, diesel, electricity, transportation activities

3. Quantities and units: numerical values paired with unit expressions (kg, L, kWh, ton-km)

The model standardizes extracted numeric values and units into a consistent machine-readable format, enabling seamless integration with downstream LCA calculations. By capturing multiple textual variants (e.g., "Concrete Used: 1,100k kg" versus "Concrete Quantity = 950k kg"), the extraction pipeline ensures high adaptability across heterogeneous infrastructure documentation. Following extraction, the system routes normalized data into an integrated LCA module, which applies standardized emission factors derived from widely used databases such as Ecoinvent and U.S. EPA inventories. Emissions are computed across three analytical levels: total lifecycle emissions (kg CO_2e), annualized emissions over the project's operational lifespan, and emissions intensity (e.g., per m2 of project area). Representative emission factors such as 0.15 kg CO_2e/kg for concrete and 1.9 kg CO_2e/kg for steel are used to translate material inventories into quantified environmental impacts. The methodology then maps calculated emissions directly to Envision Credit CR1.2, comparing project-specific lifecycle emissions against the defined baseline. Based on the magnitude of GHG reduction achieved, projects are assigned to the corresponding Envision performance tiers, "No Improvement," "Improved," "Enhanced," "Superior," or "Restorative," each linked to a defined point allocation. This mapping bridges the analytical outputs of LCA with the structured credit requirements of the Envision framework.

To operationalize the methodology, a Streamlit-based application was developed to deliver an end-to-end compliance workflow. The interface supports PDF uploads, AI-driven entity extraction, user validation steps, automated emissions calculation, and visualization of quantitative results. Supporting visualizations, including pie charts, bar graphs, and color-coded scoring tables, enhance interpretability and decision support for practitioners engaged in sustainability assessment. System validation was conducted by comparing AI-extracted attributes with manually annotated benchmarks from the case study. Extraction accuracy was quantified using precision and recall, while LCA results were benchmarked against independently performed manual calculations. The proposed compliance-checking system automates sustainability evaluation for bridge projects by integrating a custom NER model with LCA calculations and Envision Credit CR1.2 requirements. Figure 1 shows the proposed algorithm approach. Project documentation, provided as unstructured PDF files, is processed using natural language processing techniques to extract key entities, including material quantities, units, and project metadata. Extracted values are then normalized and linked to standardized emission factors to compute both lifecycle and annualized GHG emissions. These emissions are compared against Envision's baseline values to automatically classify project performance according to Envision's CR1.2 scoring tiers. The system also updates an interactive compliance checklist, providing color-coded feedback to support transparent and traceable decision-making. Implemented in Python and deployed through a Streamlit interface, the system eliminates manual data entry, enhances calculation accuracy, and demonstrates a scalable AI-driven approach for automating environmental compliance verification within the Envision framework.

Figure 1. Proposed Algorithm Approach



2.1 Life-Cycle Assessment

LCA is a method to evaluate the total environmental impact of a product, process, or project from its beginning to end, often referred to as "cradle to grave." The method helps you understand how much environmental damage a given system causes throughout its lifespan. Using the LCA for this case study reference, we aimed to evaluate how much of the project reduces GHG emissions over its operational life, compared to previous systems or alternatives. The goal of the LCA was to evaluate the environmental impacts, specifically GHG emissions, over the full 100-year life cycle of a bridge project using a functional unit (FU) of 1 m² of bridge deck over 100 years, which is a normalized unit used to compare similar structures or designs. We used five steps based on the method by Dequidt (2012).

2.1.1 System boundaries

In step 1, we defined system boundaries and life stages. The entire project was broken down into five major life-cycle phases: production, transport, construction, maintenance, and end-of-life. The production phase includes raw materials such as concrete, rebar, steel, and prefabricated parts. The transport phase considers the distance and method of delivering materials to a site. The construction phase calculates machinery usage, fuel consumption, and site work. The maintenance phase includes lighting energy, traffic emissions, and periodic repairs. Finally, the end-of-life phase includes demolition, waste transport, and recycling impacts. These five phases were explicitly defined while GHG emissions were computed separately for each phase.

2.1.2 Input data

In step 2, all quantities of materials and processes are gathered as shown in the example inputs for inventory data collection in Table 1. Input data include quantities of materials, transport distances, fuel use, electricity, and traffic activity. This raw data is later combined with emission factors (kg CO_2e per unit) to calculate total emissions.

Table 1. Example of Inputs for Inventory Data Collection

Material/Process	**Quantity**	**Units**	**Source**
Reinforced concrete	X	m^3	Contractor
Steel reinforcement	Y	tons	Contractor
Transport distance	Z	km	Estimation
Diesel fuel for machines	A	liters	Site data
Lighting consumption	B	kWh/year	Technical specification
Traffic emissions	Calculated from AADT	Vehicles/day	Road authority

2.1.3 Calculation of GHG quantities

In step 3, all GHG quantities are calculated using Arda LCA, Ecoinvent, and ReCiPe. Arda LCA was used for modeling inputs across phases, Ecoinvent supplied standard emission factors (e.g., for concrete, steel, diesel), and ReCiPe converted raw emissions (CH_4, CO, CO_2) into a common metric: kg CO_2e. Emission factors come from databases such as Ecoinvent and include concrete (~0.15–0.20 kg CO_2e/kg), steel (~1.6–2.0 kg CO_2e/kg), diesel (~2.67 kg CO_2e/liter), and electricity (depends on the regional energy mix). Each input is multiplied by its emission factor. For example, 10,000 kg of steel × 2.0 kg CO_2e/kg = 20,000 kg CO_2e, and 1,000 liters of diesel × 2.67 kg CO_2e/l = 2,670 kg CO_2e. The quantities are summed for the total emissions for all five life-cycle stages and divided by the bridge area. For example, total life-cycle emissions over 100 years are 6,665 kg CO_2e per m^2 of deck for the bridge project.

In this step, the annualized emissions are expressed per year by simply dividing the total emissions by the operational life. For example, annual emissions per square meter = total emissions per square meter ÷ building lifespan (years) = 6,665 kg CO_2e ÷ 100 years = 66.65 kg CO_2e/m^2/year. The result means that the bridge emits about 66.7 kg CO_2e per square meter per year when considering traffic use. The emissions drop to about 13.6 kg CO_2e/m^2/year if traffic use is excluded. Units used in the LCA method are kg CO_2e, kg CO_2e/year, and kg CO_2e/m^2/year. kg CO_2e, meaning kilograms of carbon dioxide equivalent, is a standard unit for measuring carbon footprints that accounts for all greenhouse gases (CO_2, CH_4, N_2O, etc.) as if they were CO_2. kg CO_2e/year is the total emissions released by the project each year over its operational lifetime. kg CO_2e/m^2/year is the yearly emissions normalized per square meter of bridge surface, which helps compare emissions between different-sized projects. The breakdown of GHG emissions by stage in terms of the percentage of total emissions is shown in Table 2.

Table 2. Breakdown of GHG Emissions by Stage (% of Total)

Life-Cycle Phase	**% of Total GHG**
Material production	~60–70%
Transport	~10–15%
Construction	~10%
Operation (lighting, traffic)	~15–20%
End-of-Life	<5%

3. Illustrative Usability Example of AI Model for Compliance-Checking System

An Envision-focused web application was developed to provide a structured and interactive workflow integrating AI-driven data extraction, lifecycle carbon assessment, and automated compliance tracking within a unified platform. The interface introduces users to the 5 Envision credit categories, displayed as visually distinct and intuitive navigation elements. Selecting a category directs users to the corresponding sub-credit, such as CR1.2 under the *Climate & Resilience* domain. Upon selecting CR1.2, the system launches the LCA module, the core computational engine of the application.

Users then upload project documentation in PDF format, regardless of layout or formatting style. Instead of requiring manual data entry, the system employs a custom-trained NER model to automatically extract structured information from unstructured text. The information includes project metadata (e.g., project name, location, operational lifespan, baseline GHG emissions) as well as detailed material and activity quantities (e.g., concrete [kg], steel [kg], diesel [L], electricity [kWh], and transport [ton-km]). The model identifies both numeric values and their corresponding units, enabling consistent data normalization before analysis.

Following extraction, the system computes total lifecycle GHG emissions, annualized emissions normalized by operational years, and emissions intensity (e.g., per square meter of bridge deck surface). These values are benchmarked against baseline emissions to quantify reduction performance. Based on the computed reduction percentage, the application automatically assigns the appropriate Envision performance level from No Improvement to Restorative and determines the associated credit points. Numerical results are supported by visual outputs such as pie charts and bar graphs illustrating emissions distribution across materials and activities, helping users identify carbon-intensive components. Beyond credit-specific analytics, the application updates a dynamic assessment checklist that aggregates the status of all sub-credits within the Climate and Resilience category. Completed evaluations, such as CR1.2, are automatically reflected in the checklist with color-coded indicators that enhance legibility and allow stakeholders to track compliance progress across the domain in real time. The workflow transforms a traditionally manual, labor-intensive evaluation process into a fully automated, transparent, and scalable system. With AI-based PDF parsing, lifecycle carbon computation, and interactive visualization, the platform enables consistent, efficient, and minimally supervised sustainability performance assessments aligned with Envision verification requirements.

Figure 2 shows the Envision sustainability framework webpage.

Figure 2. Envision Sustainability Framework

Figure 3 shows the sub-credit categories once users click the Climate and Resilience category.

Figure 3. Structure of Sub-credit Categories within the Climate
and Resilience Category

Figure 4 shows an assessment overview table once users click the GHG emission reduction credit, in which a project can achieve a maximum of 26 points if they meet all the requirements under the levels of achievement.

Figure 4. Assessment Overview Table

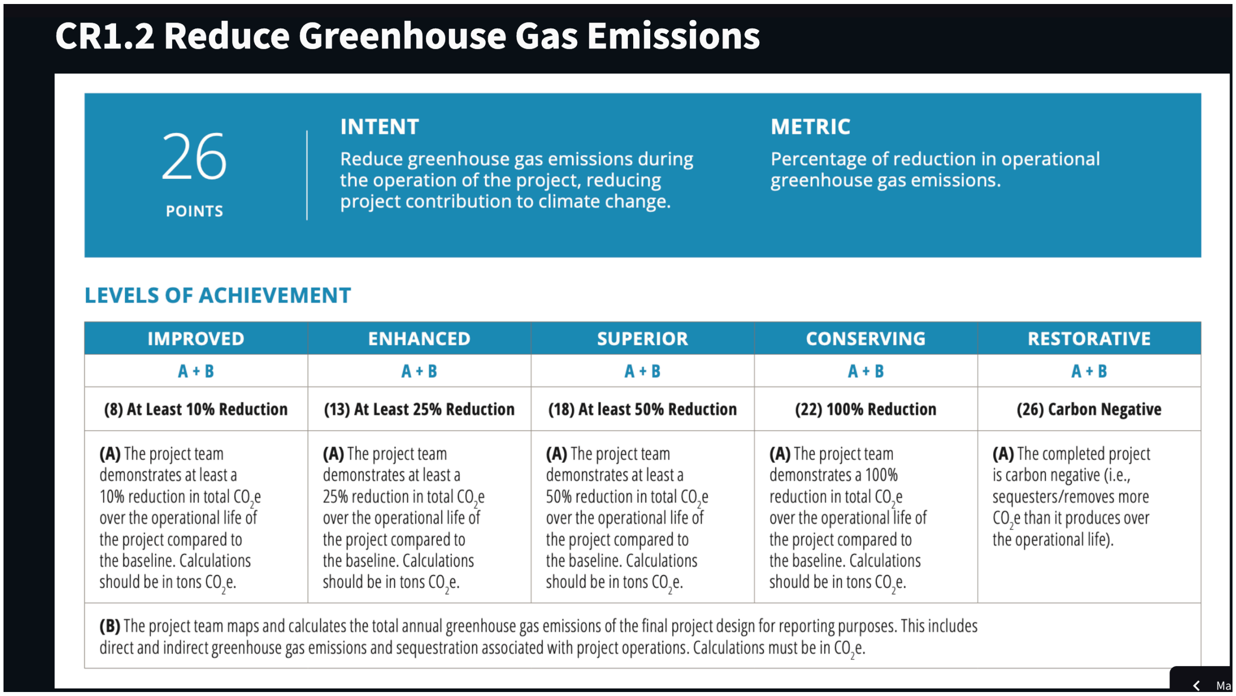
📋 Show Assessment Checklist Table

📝 Assessment Overview Table

	Credit	Subcredit	Status	Points	Max Points
0	CR1.1	Reduce Net Embodied Carbon	Not Assessed	-	20
1	CR1.2	Reduce Greenhouse Gas Emissions	Not Assessed	-	26
2	CR1.3	Reduce Air Pollutant Emissions	Not Assessed	-	18
3	CR2.1	Avoid Unsuitable Development	Not Assessed	-	16
4	CR2.2	Assess Climate Change Vulnerability	Not Assessed	-	20
5	CR2.3	Evaluate Risk and Resilience	Not Assessed	-	26
6	CR2.4	Establish Resilience Goals and Strategies	Not Assessed	-	20
7	CR2.5	Maximize Resilience	Not Assessed	-	26
8	CR2.6	Improve Infrastructure Integration	Not Assessed	-	18

Figure 5 shows the detailed requirements for CR1.2.

Figure 5. Detailed Requirement for CR1.2 Reduce GHG Emissions



CR1.2 Reduce Greenhouse Gas Emissions

26 POINTS

INTENT
Reduce greenhouse gas emissions during the operation of the project, reducing project contribution to climate change.

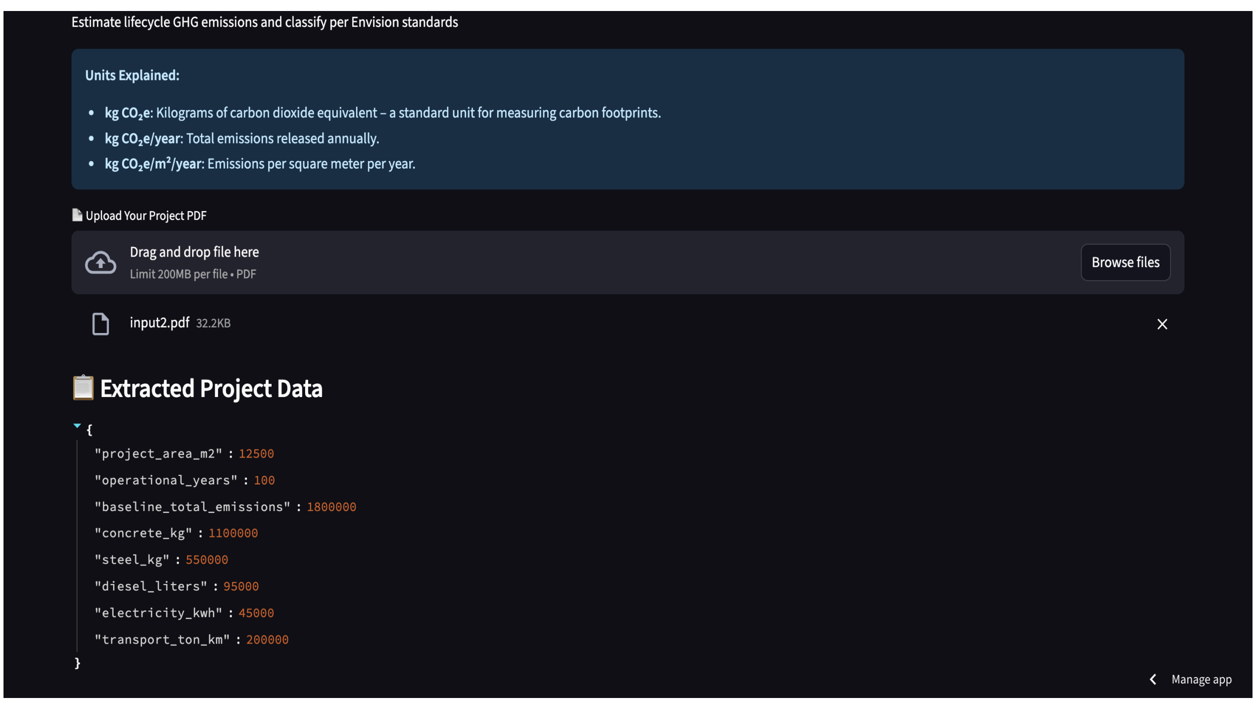
METRIC
Percentage of reduction in operational greenhouse gas emissions.

LEVELS OF ACHIEVEMENT

IMPROVED	ENHANCED	SUPERIOR	CONSERVING	RESTORATIVE
A + B	A + B	A + B	A + B	A + B
(8) At Least 10% Reduction	(13) At Least 25% Reduction	(18) At least 50% Reduction	(22) 100% Reduction	(26) Carbon Negative
(A) The project team demonstrates at least a 10% reduction in total CO_2e over the operational life of the project compared to the baseline. Calculations should be in tons CO_2e.	**(A)** The project team demonstrates at least a 25% reduction in total CO_2e over the operational life of the project compared to the baseline. Calculations should be in tons CO_2e.	**(A)** The project team demonstrates at least a 50% reduction in total CO_2e over the operational life of the project compared to the baseline. Calculations should be in tons CO_2e.	**(A)** The project team demonstrates a 100% reduction in total CO_2e over the operational life of the project compared to the baseline. Calculations should be in tons CO_2e.	**(A)** The completed project is carbon negative (i.e., sequesters/removes more CO_2e than it produces over the operational life).
(B) The project team maps and calculates the total annual greenhouse gas emissions of the final project design for reporting purposes. This includes direct and indirect greenhouse gas emissions and sequestration associated with project operations. Calculations must be in CO_2e.				

Figure 6 shows the input data and extracted project data via the AI Model for Compliance-Checking System.

Figure 6. Input and Extracted Project Data at AI Model for Compliance-Checking System

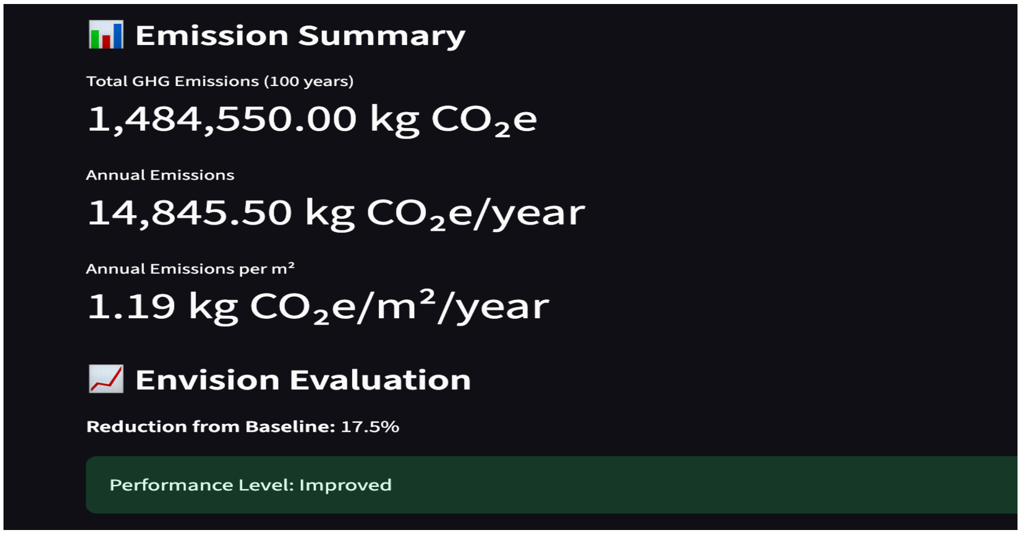


4. Experimental Results

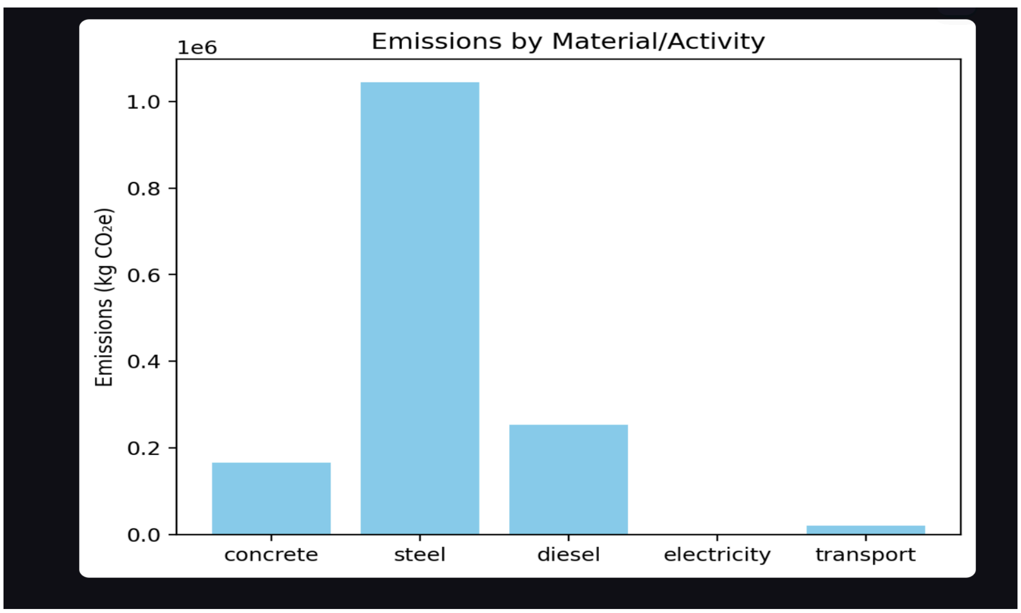
The AI-based environmental compliance checking system was evaluated using a bridge project case study, where lifecycle GHG emissions were evaluated under Envision credit CR1.2. The system successfully extracted project details such as bridge area, operational life, baseline emissions, and material quantities (concrete, steel, diesel, electricity, and transport activities) directly from uploaded PDF reports using the custom-trained NER model. These values were normalized, mapped to emission factors, and used for lifecycle emission calculations.

Figure 7 shows the results from the AI Model for Compliance-Checking System: (a) emission summary output and (b) emissions by material/activity. Figure 8 shows the results updated on the assessment overview table. It shows the emission summary dashboard generated by the system. The total lifecycle GHG emissions were calculated as 1,484,550 kg CO_2e over 100 years, translating to 14,845.5 kg CO_2e/year and 1.19 kg CO_2e/m^2/year. The Envision evaluation module classified the project's performance level as "Improved," corresponding to a 17.5% reduction from the baseline scenario. The visualizations include both bar and pie charts, showing emissions breakdown by material category, where steel (70.4%) and diesel (17.1%) were identified as the dominant contributors. In addition to emission breakdowns, the tool updates an Assessment Overview Table as shown in Figure 8. This table consolidates the performance evaluation for all climate and resilience sub-credits, with CR1.2 updated automatically upon calculation. The test case assigned a performance level of "Improved" with 5 points, highlighted in green for visual emphasis. Other credits remained marked as "Not Assessed," awaiting further AI integration for automated evaluation.

Figure 7. Output Result from AI Model for Compliance-Checking System: (a) Emission Summary Output; (b) Emission by Material/Activity



(a)



(b)

From a performance standpoint, the system demonstrated significant efficiency improvements compared to manual GHG assessments. The automated pipeline from PDF data extraction to final classification was completed in under 5 seconds, while a manual approach would require several hours of document review and calculation. Furthermore, the AI-driven NER model minimized the risk of human error by consistently extracting quantities and units, even across varying document formats.

The accuracy of the system was validated by two complementary approaches. First, we performed a manual cross-check of the extracted data and calculated emissions against the original project PDF. The results confirmed that the extracted values matched the ground truth for all tested parameters, ensuring correctness in emission calculations and Envision classification. Second, the performance of the NER model was assessed using three standard NLP metrics, Precision, Recall, and F1-score, on a hold-out validation dataset containing variations in text formats and units. As per Ameli et al. (2024), precision measures the proportion of correctly identified instances among all positive predictions, recall measures the fraction of all actual instances that are successfully detected, and the F1-score measures the harmonic mean of precision and recall.

The model achieved high performance across material and project-related entities, with an average Precision of 96.4%, Recall of 94.8%, and an F1-score of 95.6%. These results demonstrate that the NER algorithm is not only accurate but also robust against common variations in PDF formatting and terminology. The inclusion of visualizations and dynamic table updates further enhanced usability, enabling decision-makers to easily interpret project performance against Envision standards. These test results indicate that the proposed system provides both accuracy and efficiency while improving transparency and traceability of compliance checks.

Figure 8. Results Updated on Assessment Overview Table

✅ 📋 Show Assessment Checklist Table

📝 Assessment Overview Table

	Credit	Subcredit	Status	Points	Max Points
0	CR1.1	Reduce Net Embodied Carbon	Not Assessed	-	20
1	CR1.2	Reduce Greenhouse Gas Emissions	No Improvement	0	26
2	CR1.3	Reduce Air Pollutant Emissions	Not Assessed	-	18
3	CR2.1	Avoid Unsuitable Development	Not Assessed	-	16
4	CR2.2	Assess Climate Change Vulnerability	Not Assessed	-	20
5	CR2.3	Evaluate Risk and Resilience	Not Assessed	-	26
6	CR2.4	Establish Resilience Goals and Strategies	Not Assessed	-	20
7	CR2.5	Maximize Resilience	Not Assessed	-	26
8	CR2.6	Improve Infrastructure Integration	Not Assessed	-	18

5. Summary & Conclusions

This report presented the development of an AI-based environmental compliance checking system that integrates NLP with the Envision framework. It uses a custom-trained NER model to automate extraction of key project data from unstructured PDF documents and feeds them into a lifecycle carbon assessment module for classification under Envision CR1.2. The system helped automate sustainability assessments by extracting project data (materials, quantities, location, operational life, etc.) from PDFs, thereby eliminating manual data entry and calculation and integrating the extracted data into the LCA calculator to compute emissions and automate CR1.2 Reduce Greenhouse gas emissions credit evaluation. Thus, the system can help improve efficiency, accuracy, and repeatability in environmental performance scoring. The proposed approach demonstrated high extraction accuracy and consistent GHG emission calculations compared to manually validated benchmarks, while interactive dashboards and real-time assessment tables enhanced usability and transparency. The combination of automated data extraction, lifecycle emission computation, and Envision-based classification demonstrates the feasibility of integrating AI into sustainability assessment workflows for infrastructure projects.

Although the current implementation focuses on CR1.2 and performs best on structured documents, future enhancements, such as expanding support to additional Envision credits, incorporating OCR and transformer-based models, improving unit standardization, and integrating predictive modeling and human-in-the-loop validation, will strengthen adaptability and reliability. With further refinement and integration into enterprise workflows, the system holds strong potential as a scalable, intelligent tool for advancing automated sustainability assessment.

Future iterations of this research will prioritize user experience and application usability by introducing a human-in-the-loop framework, which is lacking in the current report. Specifically, upcoming development will incorporate interactive user verification screens, allowing practitioners to manually review, flag, and correct any data incorrectly extracted by the NLP module before it feeds into downstream LCA computations. For the future study, we plan to implement automated data-imputation algorithms and uncertainty-flagging mechanisms to gracefully handle partial or missing project data without stalling the overall rating computation. Addressing these real-world user scenarios will ensure the AI model is resilient, transparent, and ready for industry-wide deployment.

Appendix A

Envision App for LCA (https://envision-a.streamlit.app/)

The step-by-step procedure for using the Envision App for LCA is as follows:

1. On Envision app, click on “Open” button for “Climate and Resilience.”

2. On “Climate and Resilience” page,

    a. Scroll down click on checkbox “Show Assessment checklist table” and check for CR 1.2 in the table (It should show status as Not Assessed and points as 0).

    b. Scroll up and click on “CR 1.2 Green House Gas Emission.”

3. On “CR 1.2 Green House Gas Emission,”

    a. Scroll down, in “Upload Your Project PDF” section, upload “input2.pdf” file, once uploaded “extracted project data” will be displayed.

    b. Click on the “Submit & Calculate” button—LCA calculation will be done and results will be displayed.

    c. Now, on the left side menu, click on “Climate and Resilience” option.

4. On “Climate and Resilience” page, scroll down—click on checkbox “Show Assessment checklist table” and check for CR 1.2 in the table. (It should show status as “Improved” and points updated as “5” with green color highlight.)

Bibliography

Al Duhayyim, M., Mohamed, H. G., Alzahrani, J. S., Alabdan, R., Mousa, M., Zamani, A. S., Yaseen, I., & Alsaid, M. I. (2022). Modeling of fuzzy cognitive maps with a metaheuristics-based rainfall prediction system. *Sustainability*, *15(*1), 25.

Al Qady, M. A., & Kandil, A. (2010). Concept relation extraction from construction documents using natural language processing. *Journal of Construction Engineering and Management*, *136*(3), 294–302. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000131

Al-Ruzouq, R., Abdallah, M., Shanableh, A., Alani, S., Obaid, L., & Gibril, M. B. A. (2022). Waste to energy spatial suitability analysis using hybrid multi-criteria machine learning approach. *Environmental Science and Pollution Research*, *29*, 2613–2628.

Ameli, Z., Nesheli, S. J., & Landis, E. N. (2024). Deep learning-based steel bridge corrosion segmentation and condition rating using Mask RCNN and YOLOv8. *Infrastructures*, *9*(1), 3. https://doi.org/10.3390/infrastructures9010003

Ashayeri, M., Abbasabadi, N., Heidarinejad, M., & Stephens, B. (2021). Predicting intraurban PM2.5 concentrations using enhanced machine learning approaches and incorporating human activity patterns. *Environmental Research*, *196*, 110423.

Babu Saheer, L., Bhasy, A., Maktabdar, M., & Zarrin, J. (2022). Data-driven framework for understanding and predicting air quality in urban areas. *Frontiers in Big Data*, *5*, 822573.

Bashir Shaban, K., Kadri, A., & Rezk, E. (2016). Urban air pollution monitoring system with forecasting models. *IEEE Sensors Journal*, *16*, 2598–2606.

Cassottana, B., Biswas, P. P., Balakrishnan, S., Ng, B., Mashima, D., & Sansavini, G. (2022). Predicting resilience of interdependent urban infrastructure systems. *IEEE Access*, *10*, 116432–116442.

Cha, G.-W., Hong, W.-H., & Kim, Y.-C. (2023). Performance improvement of machine learning model using autoencoder to predict demolition waste generation rate. *Sustainability*, *15*, 3691.

Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques, and technologies: A survey on big data. *Information Sciences*, *275*, 314–347. https://doi.org/10.1016/j.ins.2014.01.015

Climate Group. (2025). *Built environment: Smarter energy use in towns and cities*. https://www.theclimategroup.org/built-environment-0

Deng, W., Mou, Y., Kashiwa, T., Escalera, S., Nagai, K., Nakayama, K., Matsuo, Y., & Prendinger, H. (2020). Vision-based pixel-level bridge structural damage detection using a link ASPP network. *Automation in Construction*, *110*, 102973.

Dequidt, T. C. E. (2012). *Life cycle assessment of a Norwegian bridge* (Master's thesis, Norwegian University of Science and Technology). Norwegian Research Information Repository. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/232313

Dorafshan, S., & Azari, H. (2020). Evaluation of bridge decks with overlays using impact echo, a deep learning approach. *Automation in Construction*, *113*, 103133.

Eken, G., Dikmen, I., & Birgonul, M. T. (2023). Using NLP for automated contract review and risk assessment. In *Proceedings of the 33rd European Safety and Reliability Conference (ESREL)* (pp. 2566–2567). https://doi.org/10.3850/978-981-18-8071-1_P030-cd

Ekici, B., Kazanasmaz, Z. T., Turrin, M., Taşgetiren, M. F., & Sariyildiz, I. S. (2021). Multi-zone optimisation of high-rise buildings using artificial intelligence for sustainable metropolises. Part 1: Background, methodology, setup, and machine learning results. *Solar Energy*, *224*, 373–389.

Energy Information Administration. (2025). *Use of energy explained: Energy use for transportation*. https://www.eia.gov/energyexplained/?page=us_energy_transportation

Fang, Q., Li, H., Luo, X., Ding, L., Rose, T. M., An, W., & Yu, Y. (2018a). A deep learning-based method for detecting non-certified work on construction sites. *Advanced Engineering Informatics*, *35*, 56–68.

Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., & Li, C. (2018b). Computer vision aided inspection on falling prevention measures for steeplejacks in an aerial environment. *Automation in Construction*, *93*, 148–164.

Fang, Q., Li, H., Luo, L., Ding, L., Luo, H., Rose, T. M., & An, W. (2018c). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Automation in Construction*, *85*, 1–9.

Fang, W., Ding, I., Zhong, B., Love, P. E. D., & Luo, H. (2018d). Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach. *Advanced Engineering Informatics*, *37*, 139–149.

Fang, W., Zhong, B., Zhao, N., Love, P. E. D., Luo, H., Xue, J., & Xu, S. (2019). A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network. *Advanced Engineering Informatics*, *39*, 170–177.

Federal Highway Administration. (2018). *INVEST in the Future with the Sustainable Highways Self-Evaluation Tool*. https://www.fhwa.dot.gov/publications/focus/13jan/13jan02.cfm

Finnveden, G., Hauschild, M. Z., Ekvall, T., Guinée, J., Heijungs, R., Hellweg, S., Koehler, A., Pennington, D., & Suh, S. (2009). Recent developments in life cycle assessment. *Journal of Environmental Management*, *91*(1), 1–21. https://doi.org/10.1016/j.jenvman.2009.06.018

Gao, X., Pishdad-Bozorgi, P., Shelden, D. R., & Hu, Y. (2019). Machine learning applications in facility life-cycle cost analysis: A review. In *Computing in Civil Engineering 2019* (pp. 267–274). https://doi.org/10.1061/9780784482445.034

Gue, I. H. V., Lopez, N. S. A., Chiu, A. S. F., Ubando, A. T., & Tan, R. R. (2022). Predicting waste management system performance from city and country attributes. *Journal of Cleaner Production*, *366*, 132951.

Guo, Y., Xu, Y., & Li, S. (2020). Dense construction vehicle detection based on orientation aware feature fusion convolutional neural network. *Automation in Construction*, *112*.

Huang, G., Ge, C., Xiong, T., Song, S., Yang, L., Liu, B., Yin, W., & Wu, C. (2021). Large scale air pollution prediction with deep convolutional networks. *Science China Information Sciences*, *64*, 192107.

Institute for Sustainable Infrastructure. (2018). *Envision: Sustainable infrastructure framework guidance manual* (3rd ed.).

Institute for Sustainable Infrastructure. (2022, July 13). *Iconic Sixth Street Viaduct in Los Angeles earns Envision Platinum*. https://sustainableinfrastructure.org/sixth-street-viaduct-seismic-improvement-project-earns-envision-platinum/

Institute for Sustainable Infrastructure. (2023). *The Next Evolution of the Envision Sustainable Infrastructure Framework*. https://sustainableinfrastructure.org/envision-v4/

Kim, J. J., & McCarthy, P. (2022). *Evaluation of sustainability determinants to develop a sustainability rating system for California infrastructure construction projects*. Mineta Transportation Institute. https://doi.org/10.31979/mti.2022.2142

Kleine Deters, J., Zalakeviciute, R., Gonzalez, M., & Rybarczyk, Y. (2017). Modeling PM2.5 urban pollution using machine learning and selected meteorological parameters. *Journal of Electrical and Computer Engineering*, *2017*, 5106045.

Koyamparambath, A., Adibi, N., Szablewski, C., Adibi, S. A., & Sonnemann, G. (2022). Implementing artificial intelligence techniques to predict environmental impacts: Case of construction products. *Sustainability*, *14*, 3699.

Lin, Y., Zhao, L., Li, H., & Sun, Y. (2018). Air quality forecasting based on cloud model granulation. *EURASIP Journal on Wireless Communications and Networking*, *2018*, 106.

Mo, X., Zhang, L., Li, H., & Qu, Z. (2019). A novel air quality early-warning system based on artificial intelligence. *International Journal of Environmental Research and Public Health*, *16*, 3505.

Naumets, S., & Lu, M. (2021). Investigation into explainable regression trees for construction engineering applications. *Journal of Construction Engineering and Management*, *147*, 04021084.

Pan, M., Yang, Y., Zheng, Z., & Pan, W. (2022). Artificial intelligence and robotics for prefabricated and modular construction: A systematic literature review. *Journal of Construction Engineering and Management*, *148*(9), 03122004.

Pan, Y., & Zhang, L. (2021). Roles of artificial intelligence in construction engineering and management: A critical review and future trends. *Automation in Construction*, *122*, 103517.

Reeder, L. (2010). *Guide to green building rating systems: Understanding LEED, Green Globes, ENERGY STAR, the National Green Building Standard, and more*. Wiley.

Renaud, J., Karam, R., Salomon, M., & Couturier, R. (2023). Deep learning and gradient boosting for urban environmental noise monitoring in smart cities. *Expert Systems with Applications*, *218*, 119568.

Sečkár, M., Schwarz, M., Golej, J., & Veverková, D. (2025). Life cycle assessment and software tools comparison. *International Journal of Environment and Sustainable Development, 24*(2), 145–162. https://doi.org/10.1504/IJESD.2025.145333

Solano Meza, J. K., Orjuela Yepes, D., Rodrigo-Ilarri, J., & Cassiraga, E. (2019). Predictive analysis of urban waste generation for the city of Bogotá, Colombia, through the implementation of decision trees-based machine learning, support vector machines and artificial neural networks. *Heliyon*, *5*(11), e02810.

Sousa, C., Teixeira, D., Carneiro, D., Nunes, D., & Novais, P. (2022). Knowledge-based decision intelligence in street lighting management. *Integrated Computer-Aided Engineering*, *29*, 189–207.

Suárez Sánchez, A., García Nieto, P. J., Iglesias-Rodríguez, F. J., & Vilán, J. A. (2013). Nonlinear air quality modeling using support vector machines in Gijón urban area (Northern Spain) at local scale. *International Journal of Nonlinear Sciences and Numerical Simulation*, *14*, 291–305.

Thimm, H. (2023). Data modeling and NLP-based scoring method to assess the relevance of environmental regulatory announcements. *Environmental Systems and Decisions, 43*, 416–432. https://doi.org/10.1007/s10669-023-09900-7

United Nations Environment Programme. (2025). *Global status report for buildings and construction (Buildings-GSR)*. UN Environment Programme and the Global Alliance for Buildings and Construction.

Wu, J., Cai, N., Chen, W., & Wang, H. (2019). Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset. *Automation in Construction*, *106*, 102894.

Wu, H., & Yao, L. (2019). Road pothole extraction and safety evaluation by integration of point cloud and images derived from mobile mapping sensors. *Advanced Engineering Informatics*, *42*, 100936.

Xu, X., Qiu, W., Li, W., Liu, X., Zhang, Z., Li, X., & Luo, D. (2022). Associations between street-view perceptions and housing prices: Subjective vs. objective measures using computer vision and machine learning techniques. *Remote Sensing*, *14*, 891.

Xu, Y., Zhou, Y., Sekula, P., & Ding, L. (2021). Machine learning in construction: From shallow to deep learning. *Developments in the Built Environment*, *6*, 100045.

Yeganeh, B., Motlagh, M. S. P., Rashidi, Y., & Kamalan, H. (2012). Prediction of CO concentrations based on a hybrid partial least square and support vector machine model. *Atmospheric Environment*, *55*, 357–365.

Yu, K. H., Zhang, Y., Li, D., Montenegro-Marin, C. E., & Kumar, P. M. (2021). Environmental planning based on reduce, reuse, recycle and recover using artificial intelligence. *Environmental Impact Assessment Review*, *86*, 106492.

Zou, H., Ge, J., Liu, R., & He, L. (2023). Feature recognition of regional architecture forms based on machine learning: A case study of architecture heritage in Hubei Province, China. *Sustainability*, *15*, 3504.

Zhang, I., Yang, X., Li, W., & Zhang, S. (2020). Automatic detection of moisture damages in asphalt pavements from GPR data with deep CNN and IRS method. *Automation in Construction*, *113*, 103119.

Zhang, J., & El-Gohary, N. M. (2011). Automated information extraction from construction-related regulatory documents for automated compliance checking. In *Proceedings of the CIB W78-W102 2011 International Conference* (pp. 574–583).

Zhang, J., & El-Gohary, N. M. (2022). Building information modeling, natural language processing, and artificial intelligence for automated compliance checking. In M. Wang, A. P. Liew, & J. C. P. Cheng (Eds.), *Research companion to building information modeling* (pp. 248–267). Edward Elgar Publishing. https://doi.org/10.4337/9781839105524.00022

Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016). Road crack detection using deep convolutional neural network. *2016 IEEE International Conference on Image Processing (ICIP)*, 3708–3712. https://doi.org/10.1109/ICIP.2016.7533052

Zhang, M., Shi, R., & Yang, Z. (2020). A critical review of vision-based occupational health and safety monitoring of construction site workers. *Safety Science*, *126*, 104658.

Zhou, Y., De, S., Ewa, G., Perera, C., & Moessner, K. (2018). Data-driven air quality characterization for urban environments: A case study. *IEEE Access*, *6*, 77996–78006.

Zhu, B., & Zhou, J. (2021). Virtual design of urban planning based on GIS big data and machine learning. *Journal of Intelligent and Fuzzy Systems*, *40*, 6263–6273.

About the Authors

Joseph J. Kim

Dr. Joseph J. Kim, PE (PI), is a Professor and Chair of the Department of Civil Engineering and Construction Engineering Management at California State University, Long Beach (CSULB). He supervised the graduate student assistant, coordinated all project activities, ensured successful project completion, and prepared the final MTI report. Dr. Kim has developed and taught three sustainability-focused courses for undergraduate and graduate students: CEM 481 Sustainability in the Built Environment, CEM 482 Sustainability in Infrastructure Systems, and CE 581 Sustainability and Green Construction. His research spans sustainability best management practices (BMPs) and the application of artificial intelligence to optimization problems in civil infrastructure systems. He has authored 115 peer-reviewed journal articles and conference papers and holds a minor in statistics, which supports his analytical research. His work has been funded by NSF TUES, Caltrans, MTI, CITT, the Metropolitan Water District, COAST, CSULB, and FDOT. Dr. Kim also played a key role in a Federal Highway Administration–funded project evaluating ITS-based pedestrian safety treatments. His responsibilities included project management, student supervision, data collection oversight, data organization, and statistical analysis using non-parametric methods. Findings from this work have been published in *Transportation Research Record* and presented at annual Transportation Research Board meetings. He has successfully completed multiple Caltrans and MTI projects and worked as a GIS specialist for the Gainesville Police Department in Florida.

Pooja D. Chavan

Pooja is a recent computer science graduate from the Department of Computer Engineering and Computer Science at CSULB who contributed to accomplishing the goals of this research project while she was a graduate student. The scope of her contributions includes assistance of the development of AI model and analysis of data. She has over three years of experience in software development, automation testing, and DevOps engineering. Her technical expertise spans building distributed systems, micro services architecture, artificial intelligence, and software quality engineering. She currently works as a Software Engineer at Daimler Truck North America, where she focuses on DevOps, process automation, and the development of scalable and efficient software solutions. She is passionate about leveraging artificial intelligence and automation technologies to solve real-world engineering and software challenges.