2026 European Conference on Computing in Construction
Corfu, Greece
July 12-15, 2026

SUPPORTING CONSTRUCTION CYBERSECURITY DECISION-MAKING THROUGH A CROWDSOURCING PLATFORM

Muammer Semih Sonkor[1], and Borja García de Soto[1]
[1]S.M.A.R.T. Construction Research Group, Division of Engineering, New York University Abu Dhabi (NYUAD), Saadiyat Island, P.O. Box 129188, Abu Dhabi, United Arab Emirates

Abstract

Construction is increasingly digitalized, and organizations are integrating operational technology, such as robots, into site operations, introducing cyber risks. This paper presents a web-based crowdsourcing platform to collect and analyze human knowledge and to provide recommendations to guide cybersecurity decision-making in the construction sector. Collected insights on risks and mitigations via open-ended questions are converted into recommendations using semantic segmentation, topic modeling, and large language models. The platform was implemented in collaboration with a UAE-based contractor on a lift-assist robot. The resulting recommendations were mapped to the foundational requirements of ISA/IEC 62443, showing alignment with six of the seven requirements in this case study.

Introduction

The construction sector has been undergoing a digital transformation, affecting site tasks. Besides the digitalization of documents and the increasing use of common data environments for data exchange and coordination, construction sites are integrating operational technology (OT), such as robotics, connected equipment, and IoT devices (García de Soto et al., 2020). While these technologies aim to improve productivity and quality, they also introduce new cyber risks into construction environments that are dynamic and often involve close interaction between humans and machines (García de Soto et al., 2020). Therefore, cybersecurity in construction is not only a concern related to confidentiality, integrity, and availability of information and resources, but also a matter of safety.

While cybersecurity emerged as a research topic in the 1970s (Lipner, 2015) and has been widely studied over time, it has not yet received enough attention in the construction domain, leaving essential areas to be explored (Salami Pargoo and Ilbeigi, 2023). This can be explained by the relatively lower maturity of digitalization in the sector (Tambwe et al., 2025) compared to more technologically advanced ones, such as energy and manufacturing. However, the increase in cyberattacks (especially ransomware) targeting construction companies (Check Point, 2025) necessitates investigating approaches suitable for the characteristics of construction projects.

Crowdsourcing is one of the effective mechanisms deployed by companies to address the lack of in-house cybersecurity expertise (Zhao et al., 2023). It has mostly been associated with bug bounty programs in the cybersecurity context, where organizations invite external security experts through an open call to identify and disclose vulnerabilities in defined systems and code bases before they are exploited by malicious actors (Malladi and Subramanian, 2020). Another common application is crowdsourced cyber threat intelligence (CTI), which involves aggregating information from open unstructured sources (e.g., social media) and structured repositories (e.g., vulnerability databases) to detect emerging threats and vulnerabilities (Sun et al., 2020).

While these approaches are highly effective in identifying isolated software vulnerabilities, they do not yield holistic, system-level guidance. Construction sites that integrate complex technologies such as robotics and connected equipment, along with building information modeling (BIM)-enabled information workflows, require deployment-oriented recommendations rather than lists of code-level findings. Such guidance is typically provided by cybersecurity consultants (often at high cost) or by in-house experts who are scarce in the construction sector, according to a recent report (Sophos, 2025). The report involved a survey of 3,700 information technology (IT) professionals from 17 countries, in which 41% of the participants from the construction and property sectors indicated that the lack of expertise was the root cause of falling victim to ransomware attacks (Sophos, 2025).

This study presents a crowdsourcing platform developed for the construction sector that gathers cybersecurity knowledge through structured open-ended questions and converts it into actionable, management-ready recommendations using a reproducible data analysis pipeline. It addresses the need for cost-effective, technology-specific cybersecurity guidance for increasingly digitized and autonomous construction sites. The platform was implemented in collaboration with a UAE-based contractor focusing on a lift-assist robot designed to support blocklaying.

In this study, the input was collected from the contractor's employees, and the resulting recommendations were mapped to ISA/IEC 62443 foundational requirements (FRs) to support systematic cybersecurity decision-making. ISA/IEC 62443 is an international cybersecurity

standard series developed for industrial automation and control systems (IACS) and broader OT environments, defining structured security requirements throughout their lifecycles.

Against this background, this study aims to address the following research questions (RQs):

- (RQ1) How should a crowdsourcing platform and analysis pipeline be structured to convert employee cybersecurity knowledge into actionable, decision-support recommendations for construction management?
- (RQ2) To what extent can internal crowdsourcing capture OT-related cybersecurity needs in construction?

RQ1 is addressed through the proposed platform architecture and the end-to-end analysis workflow presented in the Methodology section. Moreover, RQ2 is addressed by showing the extent to which the identified recommendations cover the ISA/IEC 62443 foundational requirements.

Background and Related Work

This section reviews literature identified through a targeted review of peer-reviewed studies and industry reports on construction cybersecurity and crowdsourcing.

Cybersecurity in Construction

The digital transformation of the construction industry, accelerated by the convergence of information technology (IT) and OT (e.g., autonomous machinery, IoT devices), has significantly expanded the sector's cyberattack surface (Sonkor and García de Soto, 2021). In addition to the increase in ransomware attacks targeting the sector (Check Point, 2025) in recent years, high-profile incidents such as the $25 million deepfake fraud targeting Arup (Elliott, 2025) have shown that the construction industry is on hackers' radar. The growing interest from malicious actors can be explained by the low cybersecurity preparedness among construction companies (QBE, 2024), which makes them easy targets.

To address these challenges, recent research emphasized the need for tailored vulnerability assessment methods and governance frameworks that account for the unique, dynamic nature of construction sites and their complex supply chains (Mantha et al., 2024). Moreover, Salami Pargoo and Ilbeigi (2023) indicated the need for research to develop methodologies to safeguard automated and intelligent systems used in construction projects. In addition to research, several standards and guidelines have been published to address the increasing cyber risks in the construction industry. Some examples are ISO 19650-5, which focuses on security-minded information management in BIM-enabled projects, and the code of practice "Cyber Security in the Built Environment" by the Institution of Engineering and Technology (IET). These documents provide high-level guidance to construction companies and do not focus on specific construction technologies.

Internal Crowdsourcing as a Knowledge Mobilization Mechanism

Crowdsourcing is a collaborative method for solving specific problems by leveraging a dynamic group of people formed through an open call (Pedersen et al., 2013). In cybersecurity, it is commonly utilized through programs such as bug bounties, which involve independent external experts. Although effective in many software contexts, such approaches are not always suitable for OT, where safety and availability considerations come first (Malladi and Subramanian, 2020). These concerns also apply to construction projects that use OT, such as robotic systems, IoT devices, and autonomous machinery, as they interact with the physical environment. In addition, construction systems often involve mixed ownership and responsibility (e.g., vendor-managed remote connectivity, site networks managed by the IT department, operational procedures managed by site teams), complicating external engagement.

An alternative to traditional crowdsourcing is internal crowdsourcing, where companies mobilize distributed employee knowledge for problem-solving or ideation (Ulbrich et al., 2021). One advantage of internal crowdsourcing is employees' better understanding of their organization and its operational conditions (Malhotra et al., 2017). Moreover, intellectual property and confidentiality issues are eliminated because internal participants' employee contracts already address property rights (Malhotra et al., 2017). However, companies often struggle to motivate employees to participate in internal crowdsourcing programs, and monetary rewards alone may not always be sufficient, requiring additional incentives such as recognition (Beretta et al., 2023).

The approach proposed in this study differs from bug bounties in both its participation model and intended output. Rather than expecting the participants to identify isolated code-level vulnerabilities and provide structured technical reports, it collects knowledge through open-ended questions to derive deployment-oriented, technology-specific recommendations for construction management teams. Moreover, the platform introduced in this study was not developed solely for internal crowdsourcing purposes. However, given the challenges of engaging external participants in a cybersecurity case study involving safety-critical technology, participation was limited to the contractor's employees. Therefore, the case study in this paper explores how internal crowdsourcing can be used to support cybersecurity decision-making in construction companies.

Methodology

This study adopts a design science research approach (Hevner et al., 2004), which is characterized by the creation of purposeful and innovative artifacts (in this study, a crowdsourcing platform) to solve identified practical problems.

The research was implemented through an exploratory single-case study (Yin, 2009). This methodological choice is appropriate for investigating the efficacy of new technological interventions in their real-world context, particularly in specialized and complex domains such as construction robotics.

Overview of the Crowdsourcing Platform

The proposed platform is a web-based system that enables construction companies to define a technology to be investigated, provide contextual documentation and technical details, and collect employee or external participant responses to structured open-ended questions. Submissions are then analyzed through a reproducible data analysis pipeline that identifies the common themes and derives actionable recommendations for management.

As shown in Figure 1, the high-level process flow of the proposed crowdsourcing platform involves four steps: (1) challenge creation, (2) response collection, (3) data analysis pipeline ((3a) threats and vulnerabilities and (3b) mitigations), and (4) recommendation synthesis.

System Architecture

The platform follows a standard web application structure, composed of the following main components:

- **Frontend/User Interface (UI):** Django templates (using HTML, CSS, JavaScript) implement (i) a submission dashboard for contributors to see the technical details of the investigated technology and submit free-text responses to open-ended questions, and (ii) a management dashboard for decision-makers to view the recommendations derived from the analysis.
- **Backend data handling:** Django handles HTTP requests and response processing and manages data exchange with the database.
- **Database:** MySQL stores user submissions, metadata, and analysis artifacts (e.g., derived recommendations).
- **Data Analysis Pipeline:** Python modules (separate for "threats and vulnerabilities" and "mitigations") execute the data analysis pipeline that involves data preparation, coherence-based semantic segmentation, topic modeling, and LLM-based topic labeling and summarization. The whole pipeline, including the LLM-based topic labeling and summarization, is executed locally.
- **Recommendation Synthesis:** A Python module performs LLM-based synthesis to fetch the analysis outputs from the database and convert them into a set of cybersecurity recommendations tailored to the investigated technology. The resulting recommendations are stored in the database and displayed on the management dashboard.

Overview of the Case Study

The platform was implemented in collaboration with a UAE-based contractor focusing on the MULE MZ100 (Figure 2) robot from the manufacturer Construction Robotics. The MZ100 is a lift-assist robot designed to support block laying by enabling operator-guided semi-automated lifting and positioning. The system includes an operator control panel at the end effector, load sensors intended to prevent overloads and unsafe movements, and built-in cellular connectivity that provides real-time status data and enables remote service functions (e.g., troubleshooting, diagnostics, software updates). Firmware updates require an on-site technician who connects a dedicated service cable between the machine and a laptop. These features create a mixed cyber-physical exposure profile with both physical interfaces and network-enabled functionality.

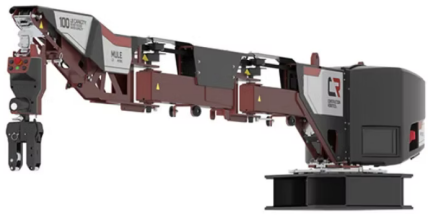

Figure 2. MULE MZ100 by Construction Robotics (Alley-Cassetty, 2026)

The case study was announced as an open call for participation within the company, targeting employees in the Design, Innovation, Digital Construction, and IT departments, as identified by the contractor. Participation was voluntary, and the analysis reported in this paper is based on responses from 11 employees who chose to contribute. The pilot ran asynchronously over two months, and participants were incentivized through the contractor's internal innovation points system managed by the Innovation department. The case study included two thematic lanes and a set of questions for each, presented on the platform: Lane A (Threats and Vulnerabilities; Q1-

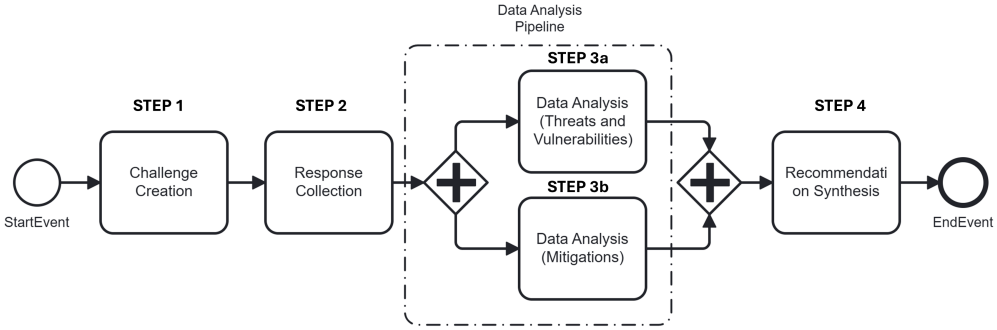

Figure 1. High-level process flow of the proposed crowdsourcing platform

Q4) and Lane B (Mitigations; Q5-Q6). All questions are presented in Table 1.

Table 1. Crowdsourcing questions used in the case study

ID	Questions
Q1	Considering the provided robotic system, what could be done to compromise the data collected? (e.g., sensor readings)
Q2	Considering the provided robotic system, what could be done to disrupt or manipulate its operation?
Q3	In your opinion, what are the most vulnerable components from the cybersecurity perspective? Why?
Q4	If you were an attacker, what would be your motivation to compromise the provided robotic system?
Q5	What would you suggest to make the system less vulnerable?
Q6	What questions should be asked to the vendor before deploying such equipment on site to identify potential cybersecurity issues?

Data Analysis Pipeline

Algorithm 1 summarizes the data analysis pipeline that converts free-text responses collected from the contributors into cybersecurity topics.

Algorithm 1. Text-to-topics data analysis pipeline

// Step 1: Fetch and Prepare the Data
1: Fetch responses from the database:
2: If Lane A: use Q1-Q4 (concatenated)
3: If Lane B: use Q5-Q6 (concatenated)
4: Clean the raw text (e.g., normalize whitespace, remove non-informative artifacts).

// Step 2: Semantic Segmentation (cosine similarity)
5: Split responses into sentences and embed them using SBERT.
6: Generate segmented corpora S_τ by applying cosine similarity thresholding to consecutive sentence embeddings over τ values (Table 2).

// Step 3: Coherence-guided selection of τ^* and k^*
7: For each candidate (τ, k) combination (Table 2), fit BERTopic on S_τ and compute c_v coherence on S_τ.
8: Select (τ^*, k^*) that maximizes c_v.

// Step 4: UMAP Hyperparameter Search (Select u^*)
9: With S_{τ^*} and k^* fixed, fit BERTopic using candidate UMAP settings (Table 2), evaluate with c_v, and select u^*.

// Step 5: HDBSCAN Hyperparameter Search (Select h^*)
10: With S_{τ^*}, k^*, and u^* fixed, fit BERTopic using candidate HDBSCAN settings (Table 2), evaluate with c_v, and select h^*.

// Step 6: Final Topic Model Fitting
11: Fit the final BERTopic model on S_{τ^*} using (k^*, u^*, h^*).
12: Extract topic keywords and representative segments for each non-outlier topic.

// Step 7: Topic Labeling and Summarization
13: Use a locally deployed LLM to generate a short label and summary for each topic (model and configuration reported in Table 2).
14: Store topic labels and summaries in the database.

The whole pipeline presented in Algorithm 1 is executed in parallel for both lanes (1) Lane A: Threats & Vulnerabilities (Q1-Q4) and (2) Lane B: Mitigations (Q5-Q6) to produce two sets of topics, including their summaries and labels. These topics are then synthesized into recommendations for management. In Algorithm 1, S_τ denotes the segmented corpus produced using threshold τ, k denotes the target topic count nr_topics, u^* denotes the selected UMAP configuration (n_neighbors, n_components) that maximizes coherence, and h^* denotes the selected HDBSCAN configuration (min_cluster_size, min_samples) that maximizes coherence.

In all search stages (i.e., Steps 3-5 in Algorithm 1), candidate configurations were evaluated using the c_v coherence metric (Röder et al., 2015) computed on the segmented corpus, and the best configuration was selected using an argmax rule. The options that produced no valid non-outlier topics were assigned a worst-case score (i.e., $-\infty$) to prevent selection. Table 2 presents the thresholds and grid-search values used in Algorithm 1.

Table 2. Values tested for each parameter in Algorithm 1

Step	Component	Parameter(s)	Values tested
Step 2	Semantic segmentation	Similarity threshold (τ)	$\{0.10, 0.15, \ldots, 0.90\}$
Step 3	Initial topic modeling search	nr_topics (k)	$\{5, 6, 7\}$
Step 4	UMAP search	n_neighbors	$\{5, 10, 15, 20, 25, 30, 35\}$
Step 4	UMAP search	n_components	$\{2, 5, 10, 15\}$
Step 5	HDBSCAN search	min_cluster_size	$\{5, 10, 15, 20, 25\}$
Step 5	HDBSCAN search	min_samples	derived from min_cluster_size (see note*)
Step 7	Topic labeling & summarization	Model; temperature	Gemma 3 27B (self-hosted via Ollama); temperature = 0

***Note:** p ∈ $\{0.10, 0.50, 1.00\}$;
min_samples = $max(1, int(p \times min\_cluster\_size))$

To identify semantic boundaries between consecutive sentences, the pipeline computes cosine similarity between their SBERT embeddings and starts a new segment when the similarity falls below the threshold τ (Steps 2 in Algorithm 1). Cosine similarity was selected since it is independent of vector magnitude, focusing strictly on the semantic orientation of the text. This choice is consistent with the original SBERT study (Reimers and Gurevych, 2019), which used cosine similarity to compare sentence embeddings. Cosine similarity between two embedding vectors (e_i) and (e_{i+1}) is calculated as follows:

$$\mathrm{sim}(\mathrm{e}_i, \mathrm{e}_{i+1}) = \frac{\mathrm{e}_i \cdot \mathrm{e}_{i+1}}{\|\mathrm{e}_i\| \|\mathrm{e}_{i+1}\|} \quad (1)$$

In Step 7 of Algorithm 1, a structured prompt template was used for LLM-based topic labeling and summarization, organized into instruction, context, input data, and output indicator, as suggested by Giray (2023). Lane-specific templates were used to maintain focus on either "Vulnerabilities and Threats" or "Mitigations". The prompt templates are not included in this paper due to page limitations; however, they can be provided upon reasonable request. Moreover, this study does not claim novelty in the individual analytical components of

Algorithm 1, since they are established techniques. Instead, the contribution lies in the proposed coherence-guided pipeline, which iteratively adjusts the semantic segmentation threshold (τ) and topic modeling configuration to maximize the coherence of the final topic-modeling output.

Recommendation Synthesis

To identify actionable recommendations for construction management, the outputs of the two analysis lanes (i.e., topic labels and summaries) were synthesized using a locally deployed LLM. The LLM (i.e., Gemma 3 27B via Ollama) was prompted to generate 5-7 actionable recommendations for construction decision-makers. The pseudocode presented in Algorithm 2 summarizes this procedure.

Algorithm 2. Recommendation synthesis workflow

1: Fetch topic labels and summaries for Lanes A and B from the database:
2: Create a synthesis prompt that instructs the model to:
- act as a Chief Information Security Officer (CISO)
- generate 5-7 actionable recommendations for a construction management team
- use only the provided topic labels and summaries as input (no other assumptions)
- write implementable and technical guidance without extra narration
3: Query the locally deployed LLM.
4: Store the resulting recommendations in the database.

This section (Methodology) addresses RQ1 by describing the proposed platform design and analysis workflow, specifying how employee knowledge was collected, processed, and converted into cybersecurity recommendations for construction management teams.

Results and Discussion

Overview of the Participants and Collected Responses

In this case study, the crowdsourcing platform received contributions from 11 employees across four departments within the collaborating construction company. Employees' departments and the top five keywords extracted from their responses (stop words removed) are presented in Table 3. Across all submissions, the Threats and Vulnerabilities lane (Q1-Q4) contained 1,183 words (108 words per employee submission, on average), whereas the Mitigations lane (Q5-Q6) contained 534 words (49 words per employee submission, on average). This difference is expected given the larger number of questions in the Threats and Vulnerabilities lane.

Platform Outputs: Actionable Recommendations for Construction Management

For traceability, Table 4 and **Error! Reference source not found.** present the topics identified by the data analysis pipeline (Figure 1) for the (i) Threats and Vulnerabilities and (ii) Mitigations lanes, respectively. These topics reflect the main themes that emerged from employees' contributions (e.g., concerns about remote access, data integrity, and physical access points). Topic summaries are not included in this paper due to page limitations; however, they are available in a repository (Sonkor and García de Soto, 2026).

Table 3. List of the contributing employees and top keywords

Employee #	Department	Top 5 Keywords
Employee 1	Design	software, data, network, communication, secure
Employee 2	Design	manipulated, opportunity, troubleshooting, telemetry, firmware
Employee 3	Design	access, program, unauthorized, hacking, mule
Employee 4	Design	remote, connectivity, connect, remotely, unsafe
Employee 5	Design	sensors, data, sensor, signals, confuse
Employee 6	Design	access, data, robot, power, control
Employee 7	Design	data, access, physical, updates, robotic
Employee 8	Design	data, equipment, connected, device, access
Employee 9	Innovation	data, security, access, incorrect, training
Employee 10	Digital Construction	access, data, machine, think, physical
Employee 11	IT	access, unauthorized, public, addresses, system

Table 4. Topics identified for threats and vulnerabilities

ID	Topic Label	Top keywords (weight)
T1	Robotic System Manipulation & Unauthorized Access	robot (0.07), sensor (0.05), weak (0.05), tampering (0.05), access (0.05), power (0.04), command (0.04), manipulation (0.04), input (0.04), feeding false (0.04)
T2	Remote Access & Physical Security Risks	access (0.07), remote (0.07), malicious (0.06), update (0.05), security (0.05), cause (0.04), control (0.04), data (0.04), physical (0.04), remote connectivity (0.03)

Table 5. Topics identified for mitigations

ID	Topic Label	Top keywords (weight)
M1	Robotic Equipment Cybersecurity: Mitigation Strategies	security (0.09), data (0.06), secure (0.06), communication (0.05), vendor (0.05), cyber (0.05), monitoring (0.05), encryption (0.04), defense (0.04), standard (0.04)
M2	Robotic Equipment Access Control & Remote Connectivity	access (0.13), remote (0.07), device (0.04), maybe (0.04), physical (0.04), robot (0.04), remote access (0.04), unauthorized (0.04), authentication (0.04), data (0.03)

The platform's recommendation synthesis step (Figure 1) then identified management-ready controls by using the identified topics as input and an LLM prompt template, ensuring that the recommendations are grounded in crowd-derived themes rather than relying on the LLM's

parametric knowledge. A list of the identified recommendations is presented in Table 6.

Table 6. Recommendations identified by the crowdsourcing platform

ID	Recommendation
R1	Implement physical access controls to the MULE MZ100, including lockable panels covering the service port and a logged, supervised access procedure for any physical interaction with the robot's internal components.
R2	Enforce multi-factor authentication (MFA) for all remote access connections to the MULE MZ100, utilizing a strong authentication method beyond passwords, such as time-based one-time passwords (TOTP) or hardware tokens.
R3	Restrict remote access to the MULE MZ100 to a pre-approved list of authorized personnel and block all incoming connections originating from public IP addresses using firewall rules.
R4	Conduct thorough vendor vetting of Construction Robotics, including a review of their security practices, compliance certifications (e.g., SOC 2), and a detailed questionnaire regarding potential vulnerabilities and incident response plans.
R5	Implement encrypted communication channels for all data transmitted to and from the MULE MZ100, utilizing TLS 1.3 or higher, and verify certificate validity regularly.
R6	Establish a logging system to record all physical and remote access attempts to the MULE MZ100, including timestamps, user IDs, and actions performed, and review these logs regularly for anomalies.
R7	Contractually obligate remote support vendors providing access to the MULE MZ100 to adhere to specific security standards and assume responsibility for any security vulnerabilities introduced through their access or actions.

The recommendations listed in Table 6 are presented as controls based on the identified topics that a construction management team can practically implement. Each recommendation suggests technical actions to safeguard the deployed technology (i.e., MULE MZ100) from different aspects (e.g., integrity, availability) and monitor it to detect potential attacks in a timely manner. Moreover, the recommendations include vendor governance actions (i.e., R4, R7), which are critical in construction OT contexts since vendors often control parts of the system lifecycle (e.g., remote support, update mechanisms) while the construction IT team might be responsible for other parts, such as deployment on site and connectivity. This might make ownership and responsibility boundaries ambiguous in projects.

Mapping Recommendations to ISA/IEC 62443 Foundational Requirements

As mentioned in the Introduction section, ISA/IEC 62443 is an international standards series that establishes cybersecurity requirements for securing IACS. IEC 62443-4-2—a part of this series—structures these requirements around seven component-level FRs (Table 7) to support holistic protection against threats. Although construction OT differs from traditional IACS, robotic equipment such as the MULE MZ100 exhibits exposure patterns comparable to those of traditional IACS (e.g., remote connectivity, maintenance interfaces) and may have operational and safety implications if compromised. In this study, the recommendations were mapped to the ISA/IEC 62443 FRs to provide a standardized way to present which security objectives are addressed, highlight potential gaps, and support interpretation. This mapping is intended to illustrate the alignment of the recommendations with the FRs and does not claim full satisfaction of any of them. The mapping is shown in Table 8.

Table 7. Foundational requirements from ISA/IEC 62443

Foundational Requirement	Description
FR1 - Identification & Authentication Control (IAC)	Ensure all users, services, and devices are uniquely identified and strongly authenticated before any system access is granted.
FR2 - Use Control (UC)	Limit authenticated entities to the minimum necessary actions via role-based authorization and enforceable policies.
FR3 - System Integrity (SI)	Preserve the integrity of software, firmware, and configurations by preventing, detecting, and recovering from unauthorized changes or code execution.
FR4 - Data Confidentiality (DC)	Safeguard data in transit and at rest through robust cryptography, sound key management, and secure handling practices.
FR5 - Restricted Data Flow (RDF)	Constrain communications by segmenting networks and enforcing least-route, allowlisted flows across defined trust boundaries.
FR6 - Timely Response to Events (TRE)	Detect, record, analyze, and act on security-relevant events promptly to contain incidents and support recovery.
FR7 - Resource Availability (RA)	Maintain essential functions during faults or attacks by resisting resource exhaustion, providing resilience/redundancy, and enabling rapid restoration.

Table 8. Mapping of recommendations to ISA/IEC 62443 FRs

Recommendation ID	Mapped Foundational Requirement(s)
R1	• FR2 - Use Control (UC)
R2	• FR1 - Identification & Authentication Control (IAC)
R3	• FR2 - Use Control (UC) • FR5 - Restricted Data Flow (RDF)
R4	• FR3 - System Integrity (SI) • FR6 - Timely Response to Events (TRE)
R5	• FR4 - Data Confidentiality (DC)
R6	• FR6 - Timely Response to Events (TRE)
R7	• FR2 - Use Control (UC) • FR3 - System Integrity (SI)

As shown in Table 8, the identified recommendations collectively address six of the seven ISA/IEC 62443 FRs

(Table 7), indicating alignment with key security objectives and thus addressing RQ2. FR7 (i.e., Resource Availability) is not addressed by any recommendations, suggesting that availability-focused controls were less prominent in employee contributions. This can be explained by both the question design and the participants' lack of cybersecurity expertise. In future case studies, this can be addressed by adding availability-oriented questions during the submission phase and involving participants with deeper knowledge of different facets of cybersecurity.

Discussion in the Construction Context

A key value of the crowdsourcing approach in this study is that it surfaced context-dependent issues that are difficult to derive from generic control catalogs alone. Although the resulting recommendations are preliminary and specific to this case study, they illustrate how employee knowledge can be translated into deployment-oriented site- and technology-specific actions for construction management. The identified recommendations can be grouped into four operational categories relevant to construction decision-making:

- **Physical access management (R1):** Construction equipment is often operated in environments where physical access is difficult to control completely. The MZ100's firmware update process involves on-site connection to a laptop via a dedicated service cable, indicating the existence of service access pathways that should be protected through physical safeguards and clear site procedures.
- **Remote access hardening (R2, R3):** Unlike most industrial facilities, construction sites mostly rely on temporary connectivity (e.g., site Wi-Fi, cellular networks), which increases the likelihood of inconsistent configurations and shared credentials. In this case study, cellular connectivity enables efficient remote software updates and troubleshooting, given the robotic system's potentially changing location as the project progresses; however, it also creates a high-impact threat surface. Safeguarding remote access through MFA and restricting it to a pre-approved list of personnel can reduce the risk during site transitions.
- **Secure communication and monitoring (R5, R6):** Construction projects generate operational data that can influence production planning and site decision-making. Therefore, telemetry and real-time status data introduce integrity concerns. Encryption and certificate validation reduce interception risk, while logging and anomaly review support detection and response.
- **Vendor governance and responsibility allocation (R4, R7):** Construction OT deployments often involve multi-party responsibility, where the vendor may handle remote support and updates, the contractor's IT team may manage connectivity, and site teams may control day-to-day operation and physical access. Due to this fragmented governance, vendor vetting and contractual obligations are essential to clearly define roles, especially when scaling OT deployment across projects and sites.

Limitations

This study has several limitations. First, the proposed platform and data analysis flow were tested on a single case study involving contributions from 11 employees. More case studies are needed to further evaluate the platform across different scenarios. Second, although an open call was made to four departments, including IT, only one IT employee responded. The remaining employees were mostly from the Design department, suggesting lower cybersecurity familiarity than IT personnel. Moreover, the responses were shorter than expected, limiting their breadth. These aspects should be kept in mind while judging the potential of the proposed methodology in this study. Finally, the mapping of the recommendations to ISA/IEC 62443 FRs (Table 8) has not yet been reviewed by external cybersecurity experts. The results should be interpreted as decision-support insights derived from employee knowledge rather than verified statements about exploitable vulnerabilities or compliance.

Conclusions and Future Work

This paper presents a web-based crowdsourcing platform that supports construction cybersecurity decision-making by gathering human knowledge through free-text contributions and converting it into recommendations for management teams. The platform was implemented in collaboration with a UAE-based contractor, and the MULE MZ100 lift-assist robot was selected as the investigated technology. A reproducible data analysis pipeline produced four topics. These topics were synthesized into seven actionable recommendations using a local LLM. The recommendations were mapped to ISA/IEC 62443 FRs to support systematic interpretation. The methodology presented in this paper demonstrates a reproducible pathway from employee inputs to decision-support recommendations, addressing RQ1. Moreover, the mapping of recommendations, with six of seven ISA/IEC 62443 FRs captured, indicates alignment with OT security objectives, addressing RQ2, and highlights Resource Availability (FR7) as a gap to target in future case studies.

Future work will focus on addressing the limitations. First, the mapping of the identified recommendations to ISA/IEC 62443 FRs will be validated by cybersecurity experts. Second, the platform will be deployed across different construction OT scenarios, and a larger participant group with more advanced cybersecurity knowledge will be targeted. Finally, the platform will be expanded to include a retrieval-augmented generation (RAG) engine that leverages relevant cybersecurity documents (e.g., standards, guidelines) to support the identified crowdsourced recommendations.

Acknowledgments

This research was partially supported by the Center for Cyber Security at New York University Abu Dhabi (CCS-AD), funded by Tamkeen under the NYUAD Research Institute Award G1104. This research was partially supported by the Center for Interacting Urban Networks

(CITIES), funded by Tamkeen under the NYUAD Research Institute Award CG001. This research was partially supported by the Center for Sand Hazards and Opportunities for Resilience, Energy, and Sustainability (SHORES), funded by Tamkeen under the NYUAD Research Institute Award CG013.

References

Alley-Cassetty, 2026. MULE MZ100. Alley-Cassetty. URL https://www.alley-cassetty.com/mulemz100 Accessed: Jan. 29, 2026.

Beretta, M., Frederiksen, L., Wallin, M., Kulikovskaja, V., 2023. Why and How Firms Implement Internal Crowdsourcing Platforms. IEEE Transactions on Engineering Management 70, 3036–3049. https://doi.org/10.1109/TEM.2020.3045118

Check Point, 2025. Global Cyber Threats September 2025. Check Point Blog. URL https://blog.checkpoint.com/security/global-cyber-threats-september-2025-attack-volumes-ease-slightly-but-genai-risks-intensify-as-ransomware-surges-46/ Accessed: Nov. 17, 2025.

Elliott, D., 2025. Cybercrime: Lessons learned from a $25m deepfake attack. World Economic Forum. URL https://www.weforum.org/stories/2025/02/deepfake-ai-cybercrime-arup/ Accessed: June 6, 2025.

García de Soto, B., Georgescu, A., Mantha, B.R.K., Turk, Ž., Maciel, A., 2020. Construction Cybersecurity and Critical Infrastructure Protection: Significance, Overlaps, and Proposed Action Plan. Preprints 2020. https://doi.org/10.20944/preprints202005.0213.v1

Giray, L., 2023. Prompt Engineering with ChatGPT: A Guide for Academic Writers. Annals of Biomedical Engineering 51, 2629–2633. https://doi.org/10.1007/s10439-023-03272-4

Hevner, A., March, S., Park, J., Ram, S., 2004. Design Science in Information Systems Research. Management Information Systems Quarterly 28.

Lipner, S.B., 2015. The Birth and Death of the Orange Book. IEEE Annals of the History of Computing 37, 19–31. https://doi.org/10.1109/MAHC.2015.27

Malhotra, A., Majchrzak, A., Kesebi, L., Looram, S., 2017. Developing Innovative Solutions Through Internal Crowdsourcing. MIT SMR.

Malladi, S.S., Subramanian, H.C., 2020. Bug Bounty Programs for Cybersecurity: Practices, Issues, and Recommendations. IEEE Software 37, 31–39. https://doi.org/10.1109/MS.2018.2880508

Mantha, B.R.K., Sonkor, M.S., García de Soto, B., 2024. Investigation of the cyber vulnerabilities of construction networks using an agent-based model. Developments in the Built Environment 18, 100452. https://doi.org/10.1016/j.dibe.2024.100452

Pedersen, J., Kocsis, D., Tripathi, A., Tarrell, A., Weerakoon, A., Tahmasbi, N., Xiong, J., Deng, W., Oh, O., de Vreede, G.-J., 2013. Conceptual Foundations of Crowdsourcing: A Review of IS Research, in: 46th Hawaii International Conference on System Sciences. pp. 579–588. https://doi.org/10.1109/HICSS.2013.143

QBE, 2024. Commercial Construction Risk Report. QBE.

Reimers, N., Gurevych, I., 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics.

Röder, M., Both, A., Hinneburg, A., 2015. Exploring the Space of Topic Coherence Measures, in: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. pp. 399–408. https://doi.org/10.1145/2684822.2685324

Salami Pargoo, N., Ilbeigi, M., 2023. A Scoping Review for Cybersecurity in the Construction Industry. Journal of Management in Engineering 39, 3122003. https://doi.org/10.1061/JMENEA.MEENG-5034

Sonkor, M.S., García de Soto, B., 2026. Data for the Construction Cybersecurity Crowdsourcing Platform. https://doi.org/10.7910/DVN/KNA5VS

Sonkor, M.S., García de Soto, B., 2021. Operational Technology on Construction Sites: A Review from the Cybersecurity Perspective. Journal of Construction Engineering and Management 147, 4021172. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002193

Sophos, 2025. The State of Ransomware 2025. Sophos.

Sun, N., Zhang, J., Gao, S., Zhang, L.Y., Camtepe, S., Xiang, Y., 2020. Data Analytics of Crowdsourced Resources for Cybersecurity Intelligence. Presented at the Lect. Notes Comput. Sci., pp. 3–21. https://doi.org/10.1007/978-3-030-65745-1_1

Tambwe, O.T., Aigbavboa, C.O., Akinradewo, O.I., Adekunle, P.A., 2025. Measures to Address Cyber-Attacks in Construction Project Data Management Processes: A Cybersecurity Perspective. IET Information Security 2025, 7398742. https://doi.org/10.1049/ise2/7398742

Ulbrich, H., Wedel, M., Dienel, H.-L., 2021. Introduction to Internal Crowdsourcing: Theoretical Foundations and Practical Applications. Springer International Publishing, Cham, pp. 1–14. https://doi.org/10.1007/978-3-030-52881-2_1

Yin, R.K., 2009. Case Study Research: Design and Methods. SAGE.

Zhao, L., Yu, X., Zhou, X., 2023. The impact of regulatory mechanisms on vulnerability disclosure behavior during crowdsourcing cybersecurity testing. Math. Biosci. Eng. 20, 19012–19039. https://doi.org/10.3934/mbe.2023841