

SPEECH-TO-UPDATE BIM: A LEAN, VOICE-ENABLED WORKFLOW FOR REAL-TIME SITE PROGRESS CAPTURE

Avi Singhal¹ & Sahil Garg²

ABSTRACT

Progress reporting on construction sites is often separated from digital model updating by several manual handling steps. Information is first noted in the field, later transcribed into reports or spreadsheets, and only afterwards reflected in BIM environments. This sequence introduces delay, repeated data handling, and opportunities for error, all of which conflict with Lean Construction principles that emphasize flow and reduction of non-value-adding activities.

This paper examines a voice-enabled reporting workflow designed to shorten this information pathway. Instead of periodic transcription, site personnel communicate progress through natural speech, allowing updates to be recorded digitally with minimal intermediate handling. The work is presented as a process intervention rather than a technology deployment. A functional prototype is used to illustrate representative use situations, and the workflow is analysed in terms of reporting cycle compression, administrative effort, and relevance to production control practices such as the Last Planner System. The findings suggest that speech-based reporting can reduce reporting effort while helping project teams maintain more up to date visibility of site progress.

KEYWORDS

Lean construction, visual management, digital twin, voice user interface (VUI), real-time monitoring.

INTRODUCTION

BIM now plays a central role in the coordination and management of construction projects; however, during execution there remains a noticeable gap between what has happened on site and what is represented in the model. In many projects, progress is recorded manually in field notebooks or spreadsheets and later transcribed into digital systems. By the time updates appear in the model, they may already be outdated, reducing their usefulness for planning and decision making (Eastman et al., 2011). In many projects, site supervisors complete reporting tasks at the end of the day, often after long working hours, which can lead to incomplete or delayed updates. This creates frustration and reduces the reliability of progress information used for planning.

From a Lean Construction perspective, this workflow contains several forms of waste:

- Waiting, as decision makers lack timely information;

¹ Post Graduate Student, Civil and Environmental Engineering Department, Indian Institute of Technology Delhi (IITD), New Delhi, India, singhalavi11@gmail.com, orcid.org/0009-0000-0702-840X

² Assistant Professor, Civil and Environmental Engineering Department, Indian Institute of Technology Delhi (IITD), New Delhi, India, gargsahil@iitd.ac.in, orcid.org/0000-0002-1303-3288

- Over processing, due to repeated handling of the same data;
- Defects, arising from transcription or interpretation errors;
- Underutilization of people, as skilled professionals spend effort on administrative reporting. (Ballard & Howell, 1998; Koskela, 1992).

Lean practice emphasizes that information should move with the work and be made available as close as possible to the time of activity (Koskela, 1992). However, construction environments are demanding and fast-paced. Field staff work in conditions where their hands are occupied, mobility is essential, and structured data entry may interrupt operational tasks (Koskela, 1992).

Speech-based interaction offers an alternative mode of reporting that aligns more closely with the realities of field work. Speech is a natural form of communication and suits the hands busy conditions typical of construction work. If speech can be converted into structured, trustworthy progress data and linked reliably to BIM elements, it has the potential to create a reporting process that is both lightweight for site users and valuable for project control. This assumption is supported by the prototype evaluation, where users were able to report progress through short spoken inputs without interrupting ongoing tasks, indicating alignment with typical site working conditions.

This paper introduces a Speech to Update BIM workflow developed with this motivation. The workflow enables a user to speak a short progress command such as “Update Beam L1B201 to fifty percent complete” which is then interpreted and mapped to the relevant BIM element. The study focuses on workflow design and practical insight gained through prototype use, viewed through a lean construction perspective.

This study contributes to Lean Construction research by (1) identifying progress reporting as a process-level source of waste, (2) proposing a speech-enabled workflow that reduces intermediate information handling, and (3) providing initial empirical evidence of reduced reporting effort through prototype-based evaluation.

Figure 1 shows the difference between traditional manual reporting workflows and the suggested speech enabled process. It also shows where Lean Construction identifies waste.

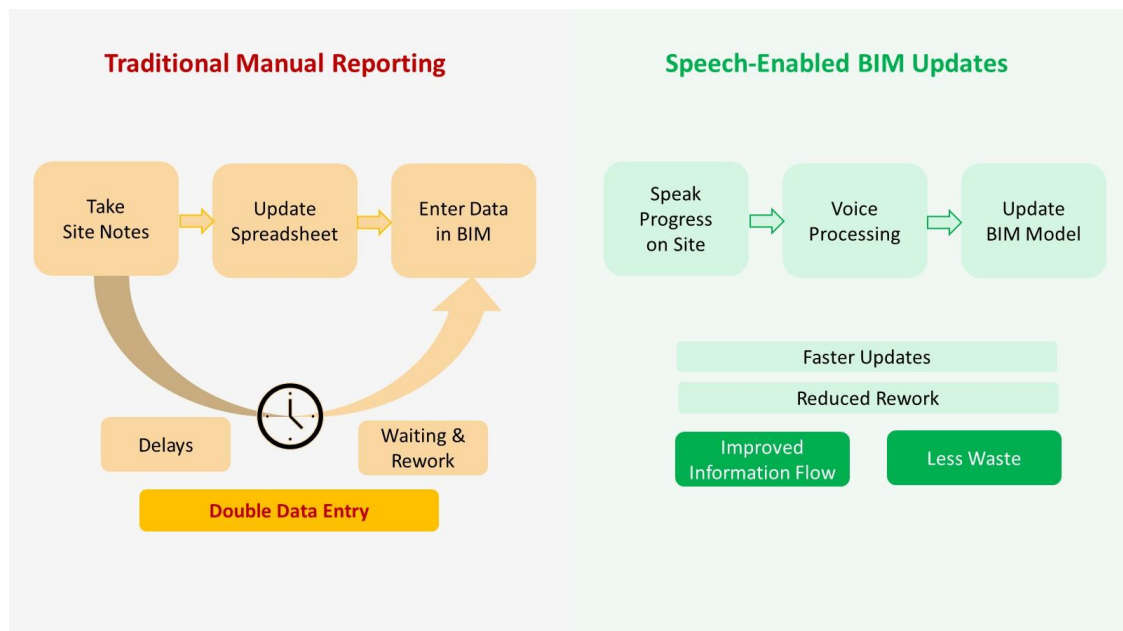


Figure 1: Problem Context and Lean Motivation (Conceptual Diagram)

BACKGROUND AND CONTEXT

Lean Construction provides a conceptual foundation for improving production flow, information reliability, and collaborative decision making across project stakeholders. Tools such as the Last Planner System (LPS) rely on accurate and timely visibility of progress in order to manage commitments, constraints, and variances (Ballard, 2000). In many implementations, however, the supporting information continues to depend on delayed, manually compiled updates (Ballard, 2000; Sacks et al., 2010).

In parallel, BIM has evolved into a central platform for integrating design, scheduling, and coordination data. Research on automated monitoring using IoT sensors, computer vision, and digital twin techniques has demonstrated strong potential, but these approaches often require additional infrastructure or operating conditions that are not universally available (Turkan et al., 2011; Golparvar Fard et al., 2012).

Voice-based and conversational interfaces represent another emerging pathway for improving human–technology interaction in construction. Existing studies have primarily focused on information retrieval or model querying rather than writing progress information back to the model (Shin & Issa, 2021). This observation creates a meaningful gap, particularly when viewed through the perspective of Lean Construction, where the priority is not just data availability, but the ease, timeliness, and effort associated with generating that data.

Recent research within the IGLC community has also emphasized the importance of improving information flow and digital integration in construction processes. Studies have shown that structured information systems and digital tracking methods can enhance production control, reduce waste, and support lean information management practices (Dave et al., 2014; von Heyl & Teizer, 2017; Ribeiro et al., 2023).

This paper responds to that gap by exploring whether speech interaction can act as a practical mechanism for reducing waste in progress reporting, while supporting continuous information flow for production control.

While Lean Construction and BIM have often been discussed as complementary but separate domains, their integration becomes particularly relevant when considering information flow. BIM provides the digital infrastructure for storing and visualizing project data, whereas Lean Construction defines how that information should move efficiently through the production system. This study positions speech-based reporting as a mechanism that bridges this gap by aligning BIM data generation with Lean principles of continuous flow and minimal handling.

RESEARCH AIM AND SCOPE

The goal of the research in this paper is to develop and demonstrate a lean motivated workflow that lets site workers use natural language speech to record progress updates and link them to BIM elements in an IFC-based environment.

The main goals of the paper are three:

- To create a modular system architecture and workflow that meets the needs of lean information flow.
- To implement and demonstrate a functional prototype that operationalizes the proposed workflow.
- To examine the practical insights and lean implications derived from representative usage scenarios.

RESEARCH METHOD

This study employs a design science research methodology, wherein a problem identified in construction practice is tackled through the design and demonstration of an artefact. The research comprised (1) pinpointing inefficiencies in traditional progress reporting workflows,

(2) creating a speech enabled BIM update workflow and prototype, and (3) assessing the artefact through representative simulated usage scenarios to extract practical insights pertinent to Lean Construction (Koskela et al., 2018).

These scenarios included structural element completion updates, reinforcement activity reporting, and segment level progress logging cases representative of routine site reporting.

The evaluation was conducted using three representative reporting scenarios reflecting common site conditions, including partial structural completion, reinforcement activity reporting, and segment-level progress updates. Each scenario was executed multiple times using both structured and unstructured speech inputs to observe system behavior. The outcomes were assessed based on successful element identification, correctness of progress interpretation, and user interaction effort.

The selection of representative scenarios was guided by the need to reflect common reporting situations where delays and manual handling typically occur. The use of both structured and unstructured speech inputs was intended to evaluate whether the system could accommodate natural variations in human communication rather than relying on rigid command formats. This approach aligns with the study's objective of reducing reporting burden while maintaining practical usability in site environments. During testing, variations in phrasing, pronunciation, and moderate background noise were intentionally allowed to reflect realistic site communication conditions.

The evaluation was conducted in a controlled prototype testing setting without external participants, where the system was operated by the authors to simulate typical site reporting scenarios. A total of three representative scenarios were tested, each executed multiple times using both structured and unstructured speech inputs. The testing environment included standard indoor conditions with moderate background noise introduced to approximate site-like communication conditions. The input consisted of short spoken progress commands representing common reporting situations. Outcomes were recorded based on successful interpretation of commands, correct mapping to BIM elements, and the time required to complete each reporting task. These observations were used to assess system behavior, usability, and reporting efficiency.

TRADITIONAL PROGRESS REPORTING AS A LEAN PROCESS PROBLEM

In many projects, progress information does not move directly from the site to digital systems. Instead, it passes through several handling stages. A supervisor observes work, makes a mental or written note, later transfers this information into a spreadsheet or report, and only after further processing does the update appear in the BIM model. Although familiar, this reporting pathway introduces delay and repeated handling of the same data.

From a lean perspective, this structure embeds several forms of waste. Time is lost while information waits to be digitised, additional effort is spent re-entering data, and errors can be introduced during interpretation or transcription. At the same time, skilled engineers devote attention to administrative reporting rather than coordination and problem solving. As a result, the model often reflects yesterday's conditions rather than the current state of work, limiting its usefulness for production control.

Figure 2 illustrates this conventional reporting pathway as a simple process map, showing where delays and non-value-adding activities occur.

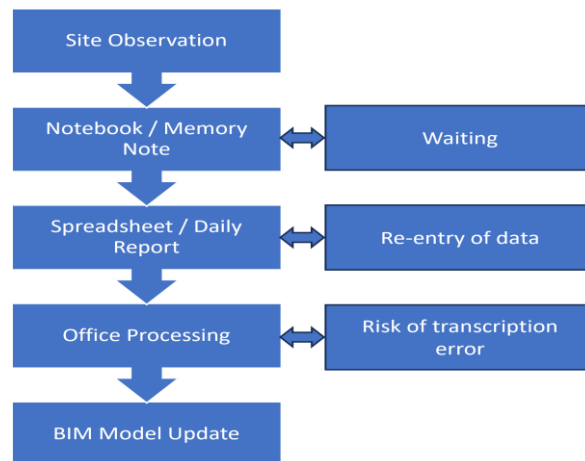


Figure 2: Conventional construction progress reporting pathway showing multiple handling stages and points of information delay.

VOICE-ENABLED WORKFLOW INTERVENTION

The proposed approach is framed as a workflow intervention rather than a technology implementation. The speech interface serves as a mechanism to restructure the information pathway, reducing the number of handling stages between site observation and digital representation.

SPEECH CAPTURE AT THE POINT OF WORK

The workflow begins with a site user (e.g., supervisor, engineer, or foreman) issuing a progress update through a mobile or web-based interface. The interaction is designed to be lightweight and minimally disruptive to site operations. Users may report task completion states such as completed, partially completed, or confirmed. The objective is to enable users to communicate in a structured manner without requiring rigid command syntax.

SPEECH TO TEXT PROCESSING

The recorded audio is passed to a speech to text engine. The chosen model emphasizes robustness to accents and moderate background noise, recognizing that construction environments are rarely quiet. The output of this stage is a textual transcription that retains the meaning of the spoken command as closely as possible (Radford et al., 2022).

HYBRID NATURAL LANGUAGE INTERPRETATION

A lightweight interpretation layer converts speech into structured progress records. The interpretation process combines keyword-based extraction with rule-based matching to identify elements such as activity type, location, and progress state from the transcribed input. These extracted entities are then mapped to BIM elements using predefined identifiers. While the approach is intentionally lightweight, it enables consistent translation of speech into structured updates without requiring complex model training.

MAPPING TO BIM ELEMENTS

The system then maps the entities to elements in an IFC based BIM dataset after they have been found. A structured mapping layer connects human readable identifiers with GlobalIDs. This ensures that updates reference precise BIM elements rather than approximate descriptions. The system is set up to give feedback instead of making an update that isn't clear (buildingSMART International, 2018).

PROGRESS LOGGING AND FEEDBACK

If a valid match is found, the interpreted progress value is saved and linked to the element. The user gets a message confirming that the command was understood and the status was recorded. This feedback loop is important for both accuracy and user trust and adoption.

The architecture is organized to remain modular and transparent, allowing future integration with dashboards, visualization modules, or planning tools.

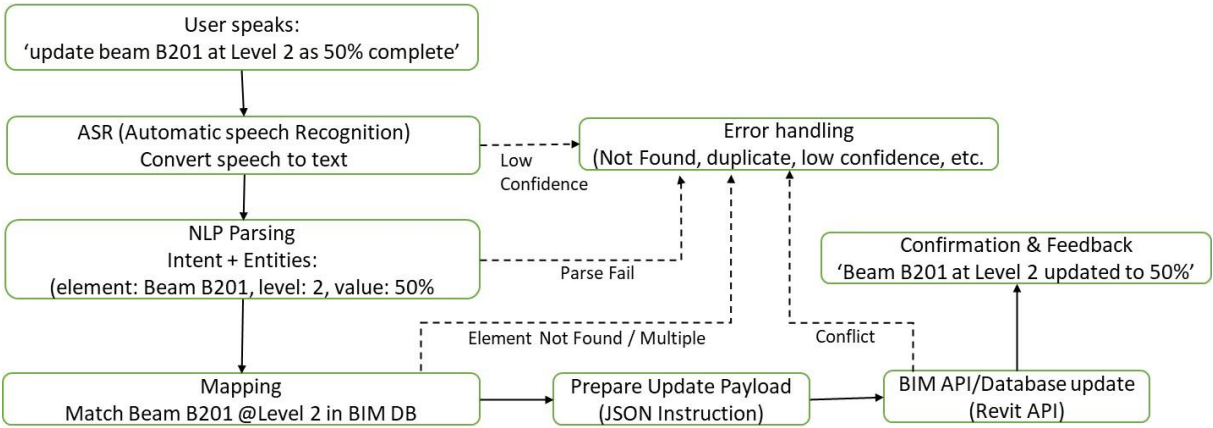


Figure 3: End to end workflow diagram (Block Flow)

Figure 3 summarizes the end-to-end flow from speech input to BIM-linked progress logging.

INFORMATION FLOW AND LATENCY REDUCTION

The main process change introduced by the voice-enabled workflow is a shortening of the reporting cycle. Instead of moving through multiple intermediate steps, progress information travels more directly from observation to digital representation. Table 1 contrasts the typical stages involved in manual reporting with those in the proposed workflow.

Table 1: Comparison of information handling stages in manual and voice-enabled progress reporting workflows.

Stage	Manual Workflow	Voice Workflow
Observation → Recording	Immediate	Immediate
Recording → Digitization	Delayed (hours/days)	Immediate digital capture
Digitization → BIM Update	Additional delay	Immediate digital capture

Table 1 presents a conceptual comparison of workflow stages based on observed process structure rather than measured time values.

In the traditional approach, the delay does not arise at the moment of observation, but in the period between recording and digitization. This lag can extend from several hours to days depending on project routines. By allowing speech input to be processed immediately, the proposed workflow reduces this gap and shifts reporting from a periodic task to a more continuous information flow. This aligns with lean thinking, which favors small, frequent information transfers over large, delayed batches (Ballard, 2000).

LEAN CONSTRUCTION PERSPECTIVE

The contribution is better understood as a lean oriented workflow change rather than purely a technical solution.

First, the workflow directly deals with waste in how information is handled. By cutting down on manual transcription and entering the same data over and over, it cuts down on administrative work that doesn't add value. Field personnel provide information immediately rather than after extended delays (Ballard & Howell, 1998; Koskela, 1992).

Second, the workflow makes it easier for information to flow. For coordination meetings, daily planning talks, and look-ahead analysis, progress data is available more quickly. When used with planning systems like LPS, this improves situational awareness and enables more responsive decision making.

Third, the method is based on a human-centered design philosophy that goes along with lean thinking. Rather than increasing reporting burden, the approach leverages natural human communication through speech to reduce interaction effort. The goal is not to take the place of professional judgement, but to make it easier to say and write down.

Finally, the workflow promotes a cultural shift towards continuous reporting instead of periodic reporting, which is very similar to lean ideas about getting feedback in small batches and moving information in small batches. The proposed workflow helps with important Last Planner System practices like more reliable look-ahead planning, daily coordination, and finding variances by cutting down on the time between site execution and progress visibility (Ballard, 2000).

LEAN IMPACT ASSESSMENT

The workflow was considered from a lean perspective to understand how the redesign affects everyday reporting practice. By removing intermediate transcription steps, the approach reduces waiting between work completion and information visibility. It also decreases over processing by avoiding repeated data handling. Because information is captured once at the source, the potential for transcription related defects is lowered. In addition, site engineers spend less time on reporting tasks and more time on coordination and supervision, supporting the lean principle of respect for people. Table 2 summarizes these impacts in relation to key lean principles.

Table 2: Observed qualitative effects of the voice-enabled workflow in relation to lean principles.

Lean Principle	Observed Effect in Reporting Practice
Flow	Continuous information movement
Waste Reduction	Removes duplicate data entry
Visual Management	Faster BIM representation
Respect for People	Reduces reporting burden

The effects summarized in Table 2 represent observed qualitative outcomes derived from prototype interaction and lean analysis, rather than directly measured quantitative metrics.

RELEVANCE TO PRODUCTION CONTROL (LPS)

More timely progress information has implications for production control. When planners have access to current site conditions, look-ahead planning can better reflect reality, emerging constraints can be identified earlier, and daily coordination discussions can be grounded in up-

to-date data. Although the prototype does not directly implement planning tools, the workflow supports the information environment on which systems such as the Last Planner System depend. In practical terms, more immediate progress updates allow weekly work planning and daily coordination discussions to reflect actual site conditions rather than delayed reports. This improves the reliability of commitments in the Last Planner System by reducing the mismatch between planned and actual work states.

PROTOTYPE DEMONSTRATION AND OBSERVED INSIGHTS

The prototype was evaluated through representative simulated scenarios that were similar to real life site reporting situations, such as when structural elements were only partially finished, when reinforcement activities were taking place, or when segment level progress updates were needed. The tests looked at both clear, structured phrases and more casual versions.

Across these simulations, several key insights emerged:

- Users were able to report progress in a natural way without having to think in database terms.
- The system was able to interpret varied ways of expressing progress without requiring rigid command formats.
- Immediate feedback increased user confidence and willingness to interact.
- How well identifiers were set up in the BIM dataset had a big effect on how well interpretation worked.

From a lean point of view, the most important thing to notice was that it took a lot less work to report progress, even in a prototype environment. A short-spoken sentence was enough to register useful data, so there was no need to fill out forms or move between screens.

The results are primarily qualitative and based on observed interactions with the prototype, supported by limited quantitative measurement as presented in Table 3. The observations reported in this section are based on direct interaction with the prototype during simulated reporting scenarios, rather than theoretical assumptions.

To provide an initial empirical indication of workflow efficiency, a simple comparative test was conducted between manual reporting and speech-based reporting. Three representative progress update tasks were performed using (1) manual spreadsheet entry and (2) voice input through the prototype. The time required to complete each reporting task was recorded. The results indicated that voice-based reporting reduced reporting time from an average of 35–45 seconds per entry to approximately 8–12 seconds, representing a reduction of nearly 70–75% in reporting effort. A summary of the observed reporting times is presented in Table 3. While limited in scope, this test provides an indicative measure of the potential time savings achievable through speech-enabled reporting. Although limited in scale, these observations provide both qualitative and indicative quantitative evidence supporting the effectiveness of the proposed workflow.

Table 3: Comparative reporting time between manual and speech-based workflows

Method	Average Time per Entry
Manual entry	35–45 seconds
Speech-based reporting	8–12 seconds

The selection of representative scenarios was guided by the need to reflect common reporting situations where delays and manual handling typically occur. The use of both structured and unstructured speech inputs was intended to evaluate whether the system could accommodate

natural variations in human communication rather than relying on rigid command formats. This approach aligns with the study's objective of reducing reporting burden while maintaining practical usability in site environments.

DISCUSSION AND RESEARCH BOUNDARIES

The implementation presented in this study is subject to defined research boundaries related to vocabulary coverage, evaluation context, and system integration. The vocabulary coverage is still limited and may not include every expression that is unique to a site. The prototype has mostly been tested in controlled settings instead of noisy, real-world settings. There has not yet been full integration with enterprise BIM platforms or role-based permissions. Also, the system doesn't yet measure productivity gains in terms of things like time saved or shorter reporting cycles.

In practical site environments, background noise and varied accents present challenges for speech-based systems. During prototype testing, it was observed that moderate background noise did not prevent successful command interpretation, although clarity of speech improved accuracy. In addition, when terms outside the predefined vocabulary or unfamiliar expressions were used, the system treated these inputs as ambiguous and prompted the user for clarification rather than applying incorrect updates. In cases where commands were ambiguous or not recognized, the system was designed to request confirmation rather than apply incorrect updates. This highlights the importance of prioritizing reliability over automation in construction environments.

These boundaries clarify the scope of the reported contribution and highlight areas where the proposed approach can be further extended and applied across broader project contexts.

As a result, the work creates a defined workflow and supporting system structure that shows how speech-based interaction can help with lean oriented progress reporting in BIM enabled projects.

STUDY CONTRIBUTIONS

This study contributes to Lean Construction and BIM enabled project management research by examining progress reporting as a process design problem rather than solely a technological challenge. It introduces a speech enabled reporting workflow that restructures how information moves from the point of work to digital models, with the aim of reducing intermediate handling stages that traditionally introduce delay and rework. The work demonstrates how voice interaction can function as a lean aligned mechanism for improving information flow, lowering administrative reporting effort, and supporting more continuous visibility of site conditions. By presenting an operational prototype and analyzing its use through representative reporting scenarios, the study provides practice-oriented insight into how natural speech can be integrated into everyday site routines without imposing rigid data entry procedures. In doing so, the paper frames voice-based progress capture as a human centered complement to existing planning and monitoring practices, highlighting its relevance to production control environments that depend on timely and reliable field information.

CONCLUSION

This paper has shown a speech enabled workflow for keeping BIM progress information up to date, and it has done so in a way that is clearly in line with the goals and philosophy of Lean Construction. The study shows that voice interaction, when well thought out, can make administrative work easier, speed up the flow of information, and make reporting practices more in line with what actually happens in the field.

Instead of seeing automation as the main goal, the contribution is to change how progress reporting is done so that it is a lean, human-centered process that technology supports but does not control. The prototype shows that it is possible to turn natural speech into structured BIM-linked progress data.

REFERENCES

- Ballard, G. (2000). THE LAST PLANNER SYSTEM OF PRODUCTION CONTROL. In *University of Birmingham Institutional Research Archive (University of Birmingham)*. <http://etheses.bham.ac.uk/4789/1/Ballard00PhD.pdf>
- Ballard, G., & Howell, G. (1998). Shielding production: essential step in production control. *Journal of Construction Engineering and Management*, 124(1), 11–17. [https://doi.org/10.1061/\(ASCE\)0733 9364\(1998\) 124:1\(11\)](https://doi.org/10.1061/(ASCE)0733 9364(1998) 124:1(11))
- Golparvar Fard, M., Peña Mora, F., & Savarese, S. (2012). Automated progress monitoring using unordered daily construction photographs and IFC Based building information models. *Journal of Computing in Civil Engineering*, 29(1). [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000205](https://doi.org/10.1061/(asce)cp.1943-5487.0000205)
- Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022, December 6). *Robust speech recognition via Large Scale Weak Supervision*. arXiv.org. <https://arxiv.org/abs/2212.04356>
- Sacks, R., Koskela, L., Dave, B. A., & Owen, R. (2010). Interaction of lean and building information modeling in construction. *Journal of Construction Engineering and Management*, 136(9), 968–980. [https://doi.org/10.1061/\(asce\)co.1943 7862.0000203](https://doi.org/10.1061/(asce)co.1943 7862.0000203)
- Turkan, Y., Bosche, F., Haas, C. T., & Haas, R. (2011). Automated progress tracking using 4D schedule and 3D sensing technologies. *Automation in Construction*, 22, 414–421. <https://doi.org/10.1016/j.autcon.2011.10.003>
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353. [https://doi.org/10.1016/s0019 9958\(65\)90241 x](https://doi.org/10.1016/s0019 9958(65)90241 x)
- buildingSMART International. (2018). *Industry Foundation Classes (IFC)*. https://www.buildingsmart.org/standards/bsi_standards/industry_foundation_classes/
- Eastman, C., Teicholz, P., Sacks, R., & Liston, K. (2011). *BIM handbook: A guide to building information modeling for owners, managers, designers, engineers, and contractors* (2nd ed.). Wiley.
- Shin, S., & Issa, R. R. A. (2021). BIMASR: Framework for Voice Based BIM Information Retrieval. *Journal of Construction Engineering and Management*, 147(10). [https://doi.org/10.1061/\(asce\)co.1943 7862.0002138](https://doi.org/10.1061/(asce)co.1943 7862.0002138)
- Dave, B., Kubler, S., Främling, K., & Koskela, L. (2014). Addressing information flow in lean production management and control in construction. In *University of Huddersfield Repository (University of Huddersfield)* (pp. 581–592). http://eprints.hud.ac.uk/id/eprint/26202/1/DaveIGLC_2014.pdf
- Von Heyl, J., & Teizer, J. (2017). Lean production controlling and tracking using digital methods. *Annual Conference of the International Group for Lean Construction*, 127–134. <https://doi.org/10.24928/2017/0238>
- Ribeiro, Y., Teixeira, J., Mêda, P., Moreira, J., Sousa, R., & Sousa, H. (2023). A Construction Information System as a lean Information Management Enabler – Case study. *Annual Conference of the International Group for Lean Construction*, 104–115. <https://doi.org/10.24928/2023/0166>