www.ricuc.cl
LinkedIn Revista Ingeniería de Construcción
ISSN: 0718-5073

Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

Chat as front end, structured data as output: A whatsapp-native AI agent for the AEC industry

Chat como front-end, datos estructurados como output: un agente de IA nativo de Whatsapp para la industria AEC

Cisterna, Diego*[1]; Weinmann, Marc; Cruz, Sebastián*****

* Co-Founder and Managing Director, Valoon GmbH, Karlsruhe, Germany.
** Research Associate, Karlsruhe Institute of Technology, Karlsruhe, Germany.
*** Co-Founder Crucerlabs, Developer, Queretaro, Mexico.

Fecha de Recepción: 16/09/2025
Fecha de Aceptación: 25/11/2025
Fecha de Publicación: 22/12/2025
PAG: 1-14

Abstract

The Architecture, Engineering, and Construction (AEC) sector continues to struggle with productivity stagnation and limited digitalization, even though daily coordination already occurs through digital channels such as email and messaging apps. This paper presents a field deployment of artificial intelligence (AI) in a large construction–mining project (~2,000 workers; ~400 machines) operating continuously, 24 hours a day and 7 days a week. The objective was to convert unstructured text, photos, and videos into structured maintenance records (tickets), aligned with a predefined schema designed to support the administration, documentation, and analysis of the interventions performed. In steady-state operation, the AI-driven system processed an average of ~77.5 maintenance tickets per day (~542.5 per week; ~3.23 per hour). After forms-based trials (both external tools and forms embedded within the messaging app) resulted in partial adoption and fragmented visibility, the in-group agent achieved near-complete capture without requiring new apps, logins, or software training. It provided in-thread transparency for acknowledgements and status updates, while cutting ~90% of manual classification and data-entry effort through automated extraction and structuring, complemented by light human-in-the-loop review. This generated significant savings by mitigating equipment downtime and its associated costs (including both rental/use costs and lost productivity). A semantics-first design (LLM + dictionaries + schema validation) produced analysis-ready data and enabled broad adoption with minimal friction. The case demonstrates a successful real-world AI deployment in AEC and highlights a transferable principle: adapt technology to the systems people already use, enabling broad adoption and frictionless data capture, and allowing intelligence to operate automatically and transparently in the background of the system.

Keywords: Artificial intelligence; large language models; messaging apps; structured data generation; AEC industry.

Resumen

El sector de Arquitectura, Ingeniería y Construcción (por sus siglas en inglés, AEC) continúa enfrentando estancamiento en la productividad y una digitalización limitada, a pesar de que la coordinación diaria ya ocurre mediante canales digitales como correos electrónicos y aplicaciones de mensajería. Este artículo presenta una implementación en terreno de inteligencia artificial (IA) en un gran proyecto de construcción de infraestructura minera (~2.000 trabajadores; ~400 unidades de maquinaria pesada) que opera de manera continua, 24 horas al día y 7 días a la semana. En este contexto, se integró un agente de IA, impulsado por un modelo de lenguaje de gran escala (por sus siglas en inglés, LLM), dentro de un grupo de WhatsApp existente utilizado para coordinar las actividades entre los operadores y mecánicos de la maquinaria del proyecto. El objetivo fue transformar texto, fotografías y videos no estructurados en registros de mantenimiento estructurados (tickets), alineados con un esquema predefinido diseñado para apoyar la administración, documentación y análisis de intervenciones realizadas. En régimen estable, el sistema basado en IA procesó en promedio ~77,5 tickets de mantenimiento por día (~542,5 por semana; ~3,23 por hora). Tras otros intentos de digitalización basados en formularios (tanto externos como dentro de la aplicación de mensajería), que lograron solo una adopción parcial y una visibilidad fragmentada, el agente integrado en el grupo alcanzó una captura prácticamente completa sin requerir nuevas aplicaciones, inicios de sesión ni capacitación en software. Proporcionó transparencia dentro del hilo para acuses de recibo y actualizaciones de estado, reduciendo ~90% del esfuerzo manual asociado a la clasificación y el ingreso de datos mediante extracción y estructuración automatizada, complementada con un proceso simple de verificación humana. Esto generó ahorros significativos al mitigar tiempos de inactividad y sus costos asociados (considerando tanto los costos de arriendo y uso de los equipos como la pérdida de productividad. Un diseño basado en semántica (LLM + diccionarios + validación de esquemas) produjo datos listos para análisis y permitió una adopción amplia con fricción mínima. Este caso demuestra una implementación exitosa de IA en la industria AEC y destaca un principio transferible: acoplar la tecnología a los sistemas que las personas ya utilizan, facilitando una adopción amplia, una captura de datos sin fricción, permitiendo que inteligencia opere de manera automática y transparente en el trasfondo del sistema.

Keywords: Inteligencia artificial; modelos de lenguaje de gran escala; aplicaciones de mensajería; generación de datos estructurados; industria AEC.

Corresponding author: diego.cisterna@valoon.chat
Co-Founder and Managing Director, Valoon GmbH, Karlsruhe, Germany.

www.ricuc.cl
LinkedIn Revista Ingeniería de Construcción
ISSN: 0718-5073
Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

1. Introduction

The Architecture, Engineering, and Construction (AEC) industry is a cornerstone of the world economy and a primary lever for infrastructure, housing, and the energy transition. Recent outlooks indicate that, despite cyclical headwinds across regions, construction remains a multi-trillion-dollar market with a medium-term growth trajectory driven by urbanization, decarbonization, and digital transformation; for example, Deloitte's 2025 Global Powers of Construction projects the global market to expand from about US$11.4 trillion in 2024 to roughly US$16.1 trillion by 2030. At the same time, country and regional forecasts show uneven momentum, with advanced economies—such as parts of Western Europe—facing near-term contractions in 2025. These dynamics heighten the sector's long-standing productivity challenge: as illustrated in Figure 1, construction's labor-productivity growth has lagged the total economy and manufacturing for decades, a gap repeatedly documented since McKinsey's 2017 analysis and reaffirmed in 2024. (Deloitte, 2025; McKinsey & Company, 2017; Yahoo Finance, 2025).

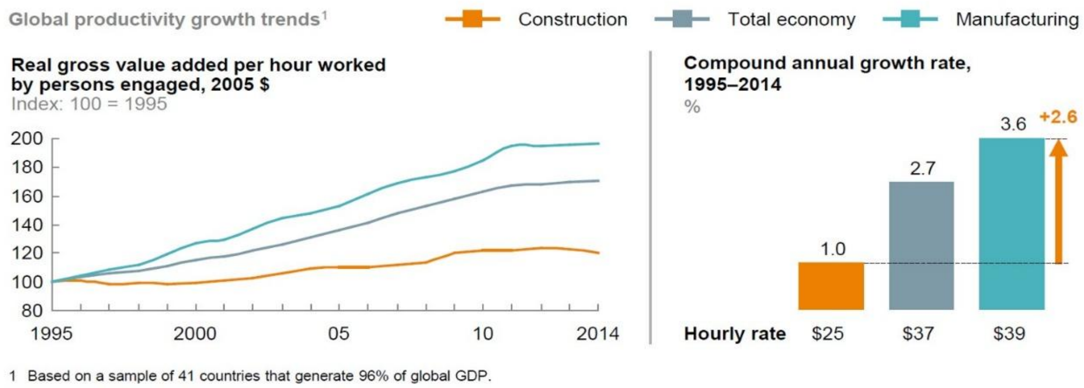


Figure 1. Global productivity growth trends (McKinsey & Company, 2017, p. 2).

Raising productivity at scale will almost certainly require deeper digitalization. Firm-level evidence from the OECD shows that digital adoption and related intangibles correlate positively with productivity growth—particularly for lagging firms—and subsequent work underscores AI's role in reducing downtime, improving coordination, and supporting data-driven decisions in production settings. Translated to AEC, the implication is clear: systematically capturing site information and turning it into computable data is a prerequisite for higher efficiency and better outcomes. (Borowiecki et al., 2021).

Yet the AEC sector has been among the least digitized of all major industries. According to the McKinsey Global Institute's Industry Digitization Index (Figure 2), construction trails far behind leading sectors across digital assets, usage, and digital labor—a structural reality that continues to explain the slow diffusion of software on job sites and the persistence of manual, paper- and chat-based workflows. Fragmentation of stakeholders, project-by-project delivery, and low process standardization compound the problem and impede data interoperability across tools. (Agarwal et al., 2016).

2. BIM and AI as drivers of digitalization in the AEC industry

The primary causes behind this are the construction industry's high degree of fragmentation and the lack of integration amongst all parties involved in the planning and building processes (Mohd Nawi et al., 2014). This results in a heterogeneous IT ecosystem of construction-related software products, as well as fragmented and disconnected knowledge silos with numerous media and information gaps (Puolitaival et al., 2018).

When it comes to construction digitization, BIM is the most prevalent term in the literature (Papadonikolaki et al., 2020). BIM enables the planning and control of projects by means of digital twins of construction buildings, and has the potential to be the backbone of digitalization in the industry (Stojanovska-Georgievska et al., 2022). The BIM method provides essential approaches for addressing the variety of software products connected to construction. Many use cases of BIM have been identified, and even their economic efficiency of implementation has been validated (Deubel, 2021). With the advent of BIM, it is feasible to digitalize additional processes, resulting in an increase in data generation.

ISSN: 0718-5073
Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

Figure 2. Industry digitization index (Agarwal et al., 2016, p. 3).

Similarly, artificial intelligence (AI) is another technology that has gained significant interest in the AEC sector in recent years. Increasing numbers of companies are developing solutions that can process data, identifying patterns in unstructured data sets, facilitating decision making, and automating processes, thereby supporting the digital transformation of the industry (Regona et al., 2022). AI is based on the analysis of massive data sets for the purpose of learning. BIM is therefore an enabler of AI applications, as it efficiently digitizes building information (Kyivska & Tsiutsiura, 2021). At the same time, AI can assist BIM in capturing, manipulating, and enabling data within its framework (Zhang et al., 2022).

Hence, a causal chain of the effects that AI deployment can have on the AEC industry can be diagrammed. As summarized in Figure 3, AI can address the AEC sector's productivity gap through three complementary routes: (i) AI–BIM synergy, where AI both supplies and processes data to accelerate BIM implementation; (ii) Automatizing processes, reducing manual capture, classification, and handoffs across site workflows; and (iii) supporting decision-making with timely, structured information (Khan et al., 2024), providing a direct path to productivity (green and blue arrows). The case study illustrated in this paper presents two real-world examples of the blue-highlighted pathways in Figure 3—Automatizing processes and supporting decision-making—clarifying how AI turns unstructured inputs into usable insights that drive information-driven decisions.

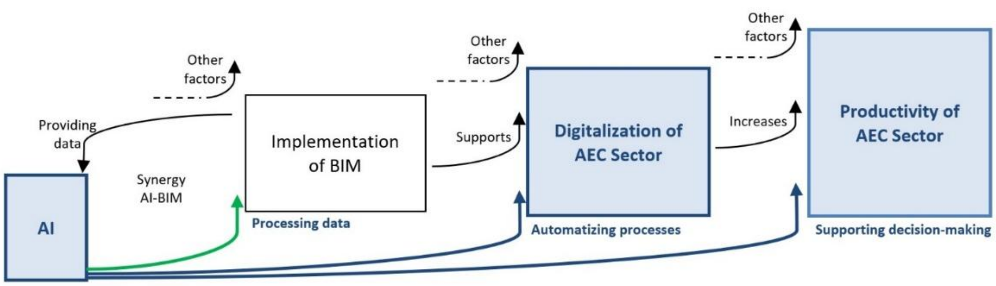


Figure 3. AI effects on productivity, digitalization, and BIM in the AEC sector.

2.1 Messaging apps as a low-friction pathway for site digitization

As the value of AI ultimately depends on timely, high-quality data, a central challenge in the AEC sector remains the capture of field information, typically exchanged informally by construction workers during daily operations. A practical way to accelerate this capture without adding change-management burden is to meet workers where they already exchange information. Messaging applications, such as WhatsApp, Telegram, or Facebook Messenger, are pervasive in site operations, and multiple studies and industry commentaries have observed their widespread use for coordination, photo exchange, and rapid issue escalation. In Germany, BauInfoConsult's Communications Monitor (2020) reports that nearly two

Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

thirds of planners and construction professionals use WhatsApp, with roughly two thirds of those users engaging daily; about three quarters employ it for internal coordination, over half for users communication, and around half for sending photos/videos and for rapid exchanges with manufacturer field staff and dealers (n=601 across architects, contractors, painters/drywallers, and HVAC installers). This pattern underscores how deeply embedded messaging already is in day-to-day site communication and documentation.

2.2 Gap, research question, and objective

Building on this conceptual pathway, we next consider where current research on AI structuring communication data still falls short. AI has demonstrated substantial value in construction management, for example, supporting the BIM–ERP integration (Álvarez et al., 2024) and through its combination with Lean Construction (Cisterna et al., 2022). However, these advances share a common characteristic: they rely on structured data environments. At the same time, recent studies highlight that unstructured data remains the dominant form of communication in construction projects, and highlight that with advancements in natural language processing (NLP) techniques, the extraction and application of this data for smart construction has become increasingly significant (Wu et al., 2022). They show that applying NLP can enhance information retrieval, improving project management efficiency and facilitating practical knowledge discovery. (Tian et al., 2021).

Yet, despite the widespread operational use of messaging platforms for digital information exchange, there is virtually no research on integrating AI directly into these native communication channels to automatically capture, classify, and structure the unstructured data already generated by project teams.

Accordingly, the research question guiding this study is: How, and to what extent, can AI automate construction workflows and strengthen data-driven decision-making, thereby contributing to increased digitalization and productivity in the AEC sector, as conceptualized in Figure 3?

To answer this question, this study examines how an AI system can turn unstructured information shared through messaging apps into structured and usable records. Using a real deployment that operates within existing WhatsApp coordination workflows, we examine how moving toward higher digitalization (by automating routine workflows and reducing manual effort) supports better decision-making through the structured information it generates, ultimately contributing to measurable productivity gains in the AEC sector.

3. Research design and methodology

This study follows a field-based single-case design, analyzing continuous operational data from a WhatsApp coordination group integrated with an AI agent. The objective is to examine how automated structuring of unstructured communication improves workflow automation and strengthens data-driven decision-making.

3.1 Data sources and observation period

The analysis draws on (i) WhatsApp group communication (>150 participants, ~700 messages/week), (ii) the structured tickets generated by the AI agent according to the maintenance schema in Table 1, and (iii) operational downtime records. Data were collected over a continuous 16-week period during steady-state operations.

3.2 Variables and metrics

From these sources, the following indicators were computed: ticket volume (daily/hourly), downtime duration, downtime by location, distribution of failure categories, adoption rates across three digitization approaches, and manual-effort reduction through automation.

3.3 Methodological workflow

The methodological sequence can be summarized as follows: (i) Ingestion of WhatsApp messages, (ii) AI parsing with schema enforcement, (iii) Normalization, (iv) Creation of a unified structured dataset, (v) Computation of the analytical indicators used in Figures 6–9.

3.4 Case study: from unstructured messaging to agentic AI–orchestrated maintenance

The examined project is a large construction–mining operation (tailings facility) with more than 2000 workers, of whom approximately 1200 are machine operators. The site runs a mixed fleet of ~400 machines (e.g., bulldozers, excavators, wheel loaders, haul trucks, personnel buses, etc.)

www.ricuc.cl
LinkedIn Revista Ingeniería de Construcción
ISSN: 0718-5073
Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

operating in four shifts, 24 hours a day, 7 days a week. Equipment requires both scheduled maintenance (e.g., oil changes, tires, engine parts) and unscheduled repairs (e.g., leaks, belt and roller failures, punctures). Two workgroups coordinate these processes:

a) Operations supervisors overseeing the ~1200 machine operators who report incidents, faults, or scheduled service needs; and

b) Mechanical crews who acknowledge receipt of equipment, perform the work, and notify when assets are ready for redeployment.

3.5 Baseline workflow and pain points

Communication was multi-channel (phone calls, WhatsApp, e-mail, and in-person exchanges). A single administrator manually consolidated messages from all these sources into an Excel tracker for both operational control and documentation/analytics. To ensure comparability over time, all entries were transcribed into a maintenance ticket schema (see Table 1).

Table 1. Maintenance ticket schema used for tracking (baseline and target).

Field	Description
Machine ID/license code	Unique equipment identifier (plate/code)
Machine type/family	e.g., bulldozer, excavator, loader, truck, bus
Reported by / Role	Supervisor/mechanic reporting party
Event type	Incident / scheduled service / other request
Description	Free-text problem or service scope
Hour-meter	Reading at report/service time
Start / End datetime	Process timestamps
Location	Area/sector of operation
Status	Reported / In workshop / In service / Closed

This workflow was inherently labour-intensive due to three compounding factors: (i) Information arrived through multiple channels that had to be reconciled; (ii) Traffic was 24/7, often outside the administrator's working hours, creating backlogs and occasional data loss; and (iii) Inputs came as unstructured natural language, with inconsistent order, spelling, and completeness—produced by operations and mechanics.

These conditions repeatedly led to missed or buried messages, delays in triage and redeployment, limited visibility of which machines were in operation versus under maintenance, and a heavy administrative burden. Despite this, every relevant post still had to be translated into the tracking fields defined in Table 1. To make this concrete, Table 2 presents anonymized excerpts from the shared WhatsApp coordination group used on site, illustrating the diversity of message styles that had to be interpreted and mapped to the structured fields in Table 1, from long narratives to terse commands and semi-structured posts.

Ideal policy vs. reality. In principle, each post should include a photo, machine code, short description, status, location, hour-meter, and request type. In practice, workers posted in varying formats and orders—often with typos or missing fields—reinforcing the need for a system that can interpret natural language and images and reliably map inputs to Table 1.

3.6 Towards zero-friction digitization with an AI agent

3.6.1 First Approach: External forms (Microsoft Forms)

To reduce manual transcription and ensure that maintenance-relevant messages could be mapped to the tracking fields in Table 1, an initial experiment was conducted with external structured forms. A link to Microsoft Forms was circulated in the group chat so that reports could be captured directly in the target schema. Despite clear instructions, adoption stabilized around ~20%, calculated as the share of users submitting reports through the forms relative to the total number of users expected to report. The reasons: leaving the group chat, loading an unfamiliar mobile UI, and completing additional steps were sufficient to deter participation.

ISSN: 0718-5073

Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

Table 2. Examples of WhatsApp messages and their LLM-based structured mapping (translated, anonymized).

Example	Original Message	Structured Mapping to Table 1
(A) Long free-text maintenance report	*"Equipment: front loader M-4172. Job: Operator seat installed (OK). Preventive maintenance: full greasing; cleaning of cabin air filters, A/C filters, and primary/secondary engine air filters. Lights checked (OK); fluid levels checked (all OK). Wear elements inspected; bucket shims due for replacement, time window: 16 to 18. Hour-meter: 4275. Status/Location: Equipment operational in [Plant Sector A1]."*	• **Machine ID:** M-4172 • **Machine type/family:** Front loader • **Event type:** Scheduled / Preventive • **Description:** Operator seat installed (OK). Preventive maintenance: full greasing; cleaning of cabin air filters, A/C filters, and primary/secondary engine air filters. Lights checked (OK); fluid levels checked (all OK). Wear elements inspected; bucket shims due for replacement. • **Hour-meter:** 4275 • **Start time:** 16:00 • **End datetime:** 18:00 • **Location:** [Plant Sector A1] • **Status:** Operational
(B) Short free-text maintenance report	*"Dear team, please return dump truck M-9089 to the workshop; it is a backup unit. Location: [Haul Road B] → [L-B3]."*	• **Machine ID:** M-9089 • **Machine type/family:** Dump truck • **Event type:** Transfer to workshop (unscheduled) • **Location:** [L-B3]
(C) Semi-structured message	*"Equipment: M-2069. Hour-meter: 2491. Service date: 02.10.2024. Work performed: Tire pressure normalization. Start/End: 03:20–03:45. Status: OPERATIONAL. Location: [Sector C2]."*	• **Machine ID:** M-2069 • **Event type:** Unscheduled repair • **Description:** Tire pressure normalization • **Hour-meter:** 2491 • **Start time:** 03:20 • **End Time:** 03:45 • **Location:** [Sector C2] • **Status:** Operational

3.6.2 Second Approach: In-channel forms (Valoon chatbot, 1-to-1)

In response to that limitation, a software tool called "Valoon" was implemented. Valoon introduced a WhatsApp-based chatbot that captures form inputs via WhatsApp Flows—interactive forms that open directly inside the WhatsApp conversation, based on the same schema used in the external forms. Once completed, the chatbot automatically converts these inputs into maintenance tickets aligned with the tracking table, ensuring that field information flows directly into the system. Adoption rose to ~60%, calculated as the share of workers registered and submitting reports through Valoon compared to the total number of workers expected to report. Nevertheless, a portion of workers continued to use the original group chat rather than interact with the bot. Two factors were consistently observed: (1) although embedded in WhatsApp, the Valoon bot still followed a slightly different interaction flow than the group's established routine, and (2) the team valued the transparency of exchanging messages in a common group (instead of a one-to-one chatbot interaction)—even at the cost of higher message volume—so that all events remained visible to the ~150 participants. As a result, acceptance improved but did not reach 100%, and communication continued through two parallel systems (the WhatsApp group and Valoon chatbot).

3.6.3 Third Approach: In-channel data capture (Valoon AI Agent)

Given these residual gaps, a further step was pursued with a zero-friction design. The aim was to keep the conversation exactly where it already occurred—in the WhatsApp group—while orchestrating the process silently. Leveraging the WhatsApp Business API, an LLM-based agent was added by Valoon directly to the existing group to operate in the background. The agent observes messages and converts unstructured inputs—text, photos, and videos—into structured records mapped to the structure of Table 1. Through connection to LLM APIs, entities (machine IDs, hour-meter values, locations, event types) are extracted and normalized against controlled dictionaries; attached images are AI-analysed to infer machine family or read visible codes.

Figure 4 provides an illustration where the system identifies both the visible machine code and the corresponding equipment family.

ISSN: 0718-5073
Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

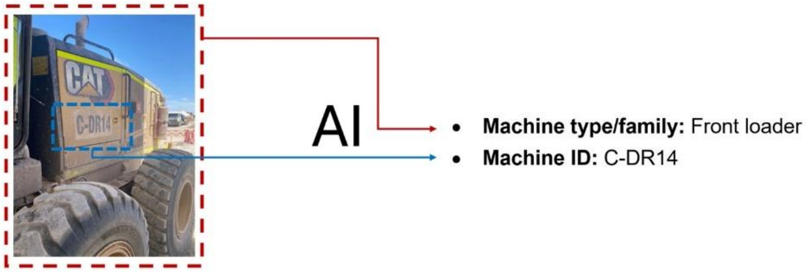


Figure 4. Example of image-based recognition of machine code and family (type of machine).

The following Table 3 summarizes the approaches and their observed outcomes, highlighting trade-offs in friction and transparency.

Table 3. Summary of digitization approaches.

Dimension	Approach 1 External forms (Microsoft Forms)	Approach 2 In-channel forms (Valoon chatbot, 1-to-1)	Approach 3 — In-channel data capture (Valoon AI Agent)
Channel/context	Outside WhatsApp	WhatsApp (private chat with bot)	WhatsApp group (existing thread)
Interaction pattern	Link → open form → submit	Chatbot form steps inside WhatsApp	Passive listener → auto-structuring; prompts only for gaps
Approx. adoption	~20%	~60%	~100% (few users report through email or telephone)
Strengths	Direct mapping to Table 1; simple to deploy	Familiar channel; entries flow directly to the tracking table	Zero change management; preserves group workflow; text+image understanding
Limitations observed	Context switch; unfamiliar UI; extra steps deter participation	3 clicks vs. 1; some users remained in group chat; preference for common-group transparency	Requires group governance and controlled dictionaries; human-in-the-loop for edge cases
Transparency (group visibility)	Not in group thread	Lower than group posts (1-to-1 bot)	High (events visible to all participants)
Result	Discontinued	Improved but parallel systems persisted (group chat + Valoon)	Retained as an operating solution; outcomes quantified later

4. Technical architecture and implementation

This section describes the design of the WhatsApp-native AI agent, focusing on how Valoon's platform features and a schema-bound parsing pipeline were used to deliver fully structured maintenance tickets from unstructured WhatsApp traffic. The system is designed for zero-friction adoption: crews continue using their existing WhatsApp group, while an AI agent runs in the background to transform messages into validated, analysis-ready records.

4.1 Valoon platform and native structures

Valoon provides a ticketing and workflow platform with a scope-based data model:

• **Scopes and Labels:** Each Scope represents a structured field (e.g., machine_id, event_type, status) and contains a defined set of Labels, which serve as enumerated values.

• **API endpoint:** Scopes and their Labels are retrieved via GET /Scopes/?expand=labels, ensuring downstream services can dynamically load field definitions and allowed values.

• **Contact management:** Users, phone numbers, and roles are maintained in Valoon's contacts database, enabling deterministic mapping of chat participants to roles (e.g., Supervisor, Mechanic).

• **Ticket model:** Tickets combine Scope-defined fields, attachments, and timestamps, making every record auditable and interoperable with analytics tools.

ISSN: 0718-5073
Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

This structure allowed the AI agent to fully align with Valoon's schema: field definitions and enumerations are authoritative, removing guesswork and post-processing.

4.2 LLM integration and schema enforcement

The parsing engine uses OpenAI's structured outputs API to guarantee JSON schema compliance (OpenAI Structured Outputs; JSON Schema).

• **Schema-first design:** The JSON Schema describing each Scope includes a description for every field and, when relevant, explanations of each allowed value. These definitions are injected into the LLM call, eliminating the need for few-shot examples.

• **Strict schema binding:** The parser does not generate free-form responses; its output is validated directly by the OpenAI structured-output mechanism, ensuring complete adherence to the schema.

• **LLM provider abstraction:** Although OpenAI's API was used in this deployment, the architecture is designed to accommodate other LLM providers or local models. Running an on-premises LLM would keep all processing within the organization's perimeter, which is operationally preferable in environments with strict data sovereignty requirements.

4.3 End-to-end data flow

Messages flow from WhatsApp to a broker and Redis-based queue, where workers invoke the schema-bound parser to create validated JSON tickets. Tickets are cached, enriched with context, and pushed to Valoon and other destinations. This design ensures deterministic output, fault tolerance, and alignment with Valoon's Scope-based schema. Figure 5 illustrates the entire pipeline, implemented as a Valoon AI Agent.

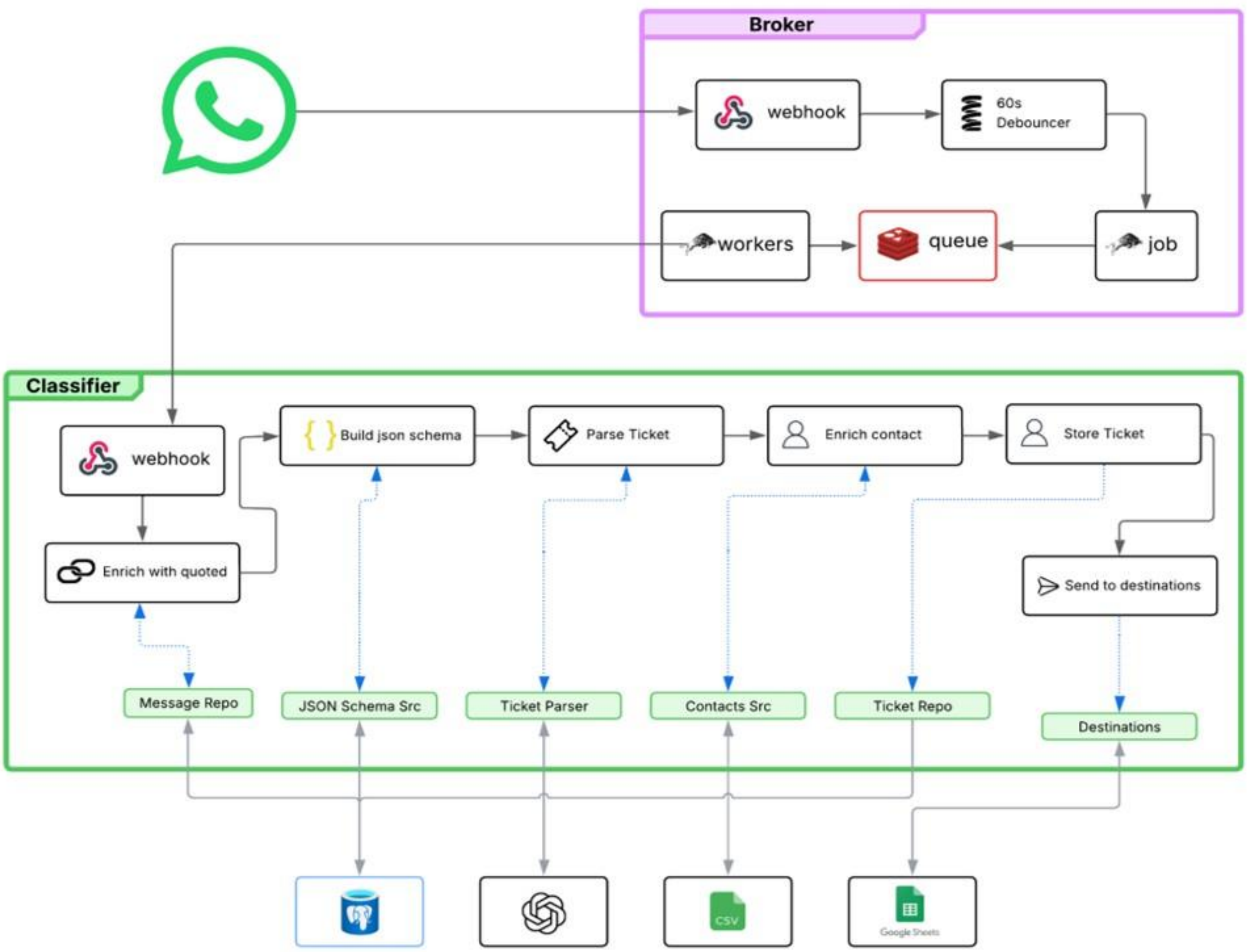


Figure 5. Valoon AI Agent pipeline.

www.ricuc.cl
LinkedIn Revista Ingeniería de Construcción
ISSN: 0718-5073
Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

1. Ingestion (WhatsApp to Broker): Messages are captured through a WhatsApp API and sent to a broker via a webhook. A 60-second debouncer groups consecutive messages from the same participant (e.g., a photo followed by a caption or correction) into a single processing unit.

2. Queueing: The broker enqueues jobs in Redis, and stateless workers pull from this queue for processing. This design enables horizontal scaling and resilience under high message volume.

3. Schema loading: Before parsing, each worker loads Scope definitions and their enumerated Labels dynamically from GET /Scopes/?expand=labels, as well as contact information from Valoon's API, ensuring strict alignment with the platform schema.

4. Parsing: The LLM-based Ticket Parser converts the grouped message bundle into a JSON object that conforms to the schema. Quoted or referenced messages are merged to preserve conversational context and link updates or closures to prior reports.

5. Enrichment: The parsed output is normalized using controlled dictionaries for machine aliases, families, and location codes, while participants are resolved to registered roles via Valoon's contact registry.

6. Persistence and fan-out: After validation, the structured ticket is stored in the Ticket Repository (PostgreSQL) and published to Valoon via POST /tickets as the authoritative record. Optional outputs, such as dashboards in Microsoft Sheets, are updated in parallel for reporting and monitoring.

4.4 Robustness, idempotency, and fault tolerance

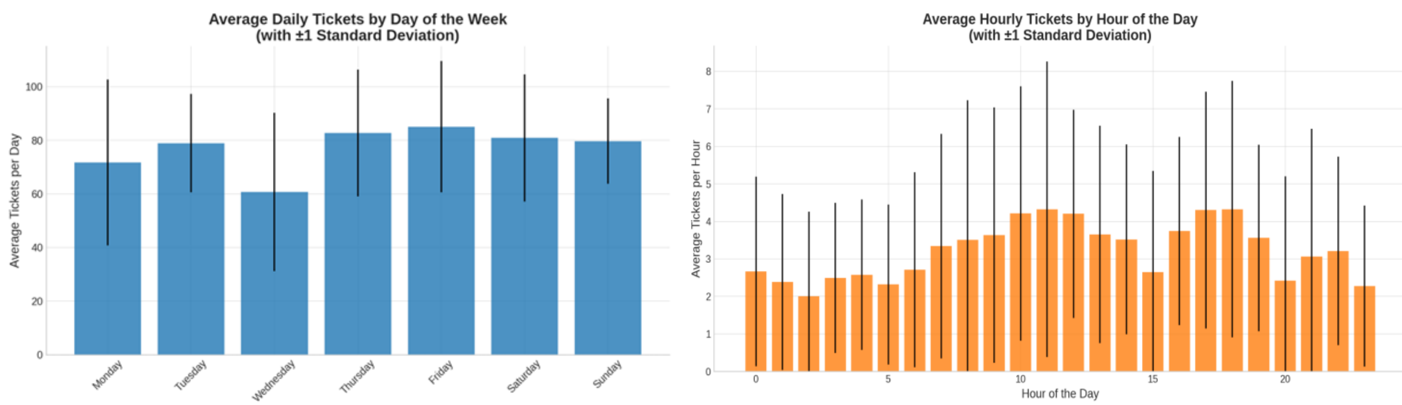
The system incorporates several features to ensure deterministic, production-grade operation:

- **Caching and idempotency:** After validation, the parsed ticket JSON is cached in the Ticket Repository. When the broker retries a job (e.g., due to a downstream failure), workers reuse this cached object, eliminating duplicate LLM calls and guaranteeing deterministic outputs.
- **Retry logic:** Failures in the LLM provider, destinations, or other integrations trigger controlled retries with no data loss.
- **Debouncing:** Multi-message posts are merged into one processing job, reducing noise and preventing duplicates.
- **Quoted message enrichment:** Replies and follow-ups are automatically linked to their source messages, maintaining conversation-level context.
- **Service isolation:** The broker, queue, parser, and storage layers are decoupled, allowing independent recovery from failures.

This architecture demonstrates a schema-first, semantics-first approach: Valoon's Scope model defines a strict data contract, while the LLM operates as a controlled parser rather than a generative agent. The result is a production-ready system capable of converting high-volume WhatsApp traffic into fully normalized tickets, with caching and retry safety ensuring reliability under operational load.

4.5 From unstructured chat to actionable analytics

By embedding the agent directly in the WhatsApp group, routine free-form exchanges were converted into structured maintenance records. What previously existed as a constant, organic flow of messages is now a dataset where patterns can be detected, and operational actions can be proposed with measurable expected outcomes. The following visualizations exemplify how raw conversation was transformed into actionable intelligence. As shown in Figure 6, average daily and hourly ticket volumes exhibit recurrent rhythms in reporting. The system reveals recurrent rhythms in reporting—higher ticket volumes toward the end of the week and clear intra-day peaks (Figure 6). These patterns support more precise planning of preventive maintenance shifts and resource allocation.

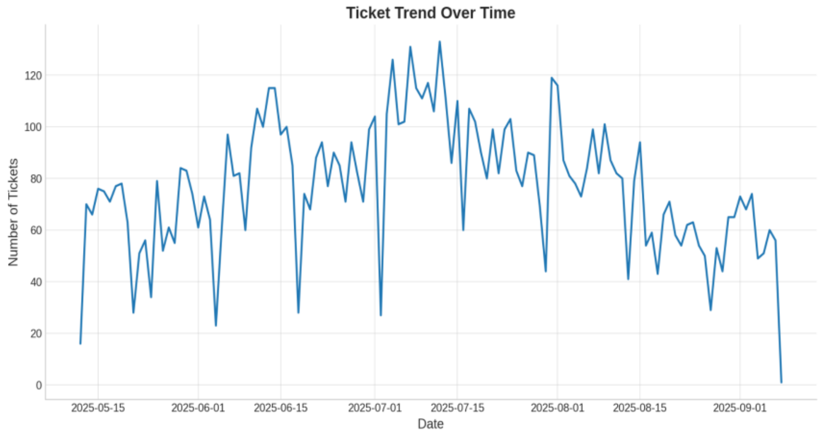


(a) (b)

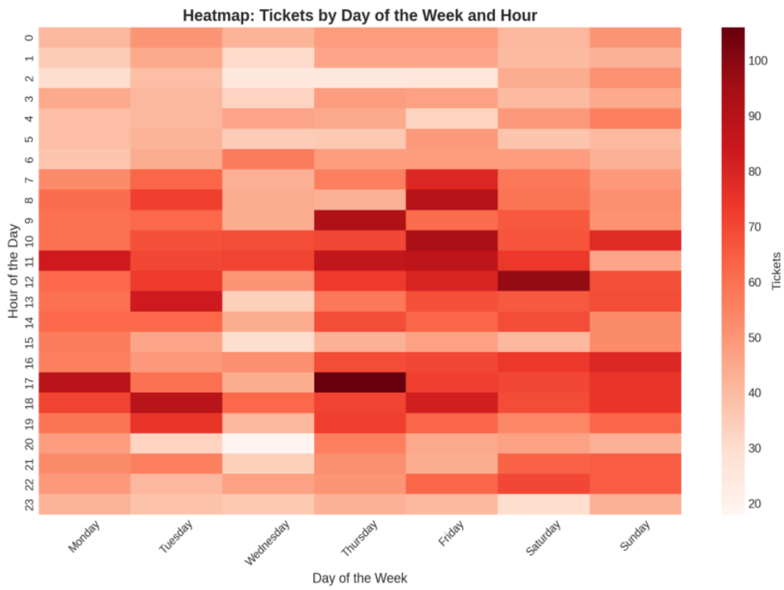
Figure 6. Average Daily and Hourly Tickets (with variability).

Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

The continuous capture over time illustrates the stability of the solution (Figure 7a). Toward the end of the observation period, the number of reports gradually decreased, reflecting the fact that the project was approaching completion and the workforce on site was being reduced. The weekday–hour heatmap (Figure 7b) highlights recurring interaction effects—not only evening clusters on Fridays but also a distinct mid-week peak on Wednesdays.



(a)



(b)

Figure 7. Ticket Trend and Heatmap.

In discussion with the users, this pattern was traced to the shift change that takes place at that time, when operators habitually pass pending tickets to the incoming crew. These findings show how normalized chat data can reveal hidden operational rhythms and provide a direct basis for aligning staffing and resource allocation with real demand windows.

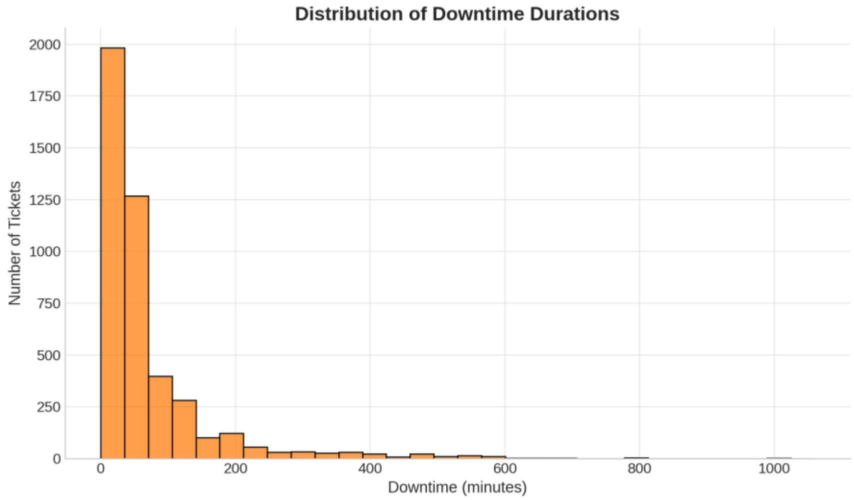
Most interventions close quickly, but the histogram and boxplot (Figure 8a) expose a long tail of extended repairs. These outliers are candidates for root-cause analysis to reduce costly, prolonged stoppages.

Averaging by site (Figure 8b) highlights certain locations (L1, L2) with systematically higher downtime. This points to either environmental conditions or local workshop capacity issues and suggests where reinforcement could yield the largest impact.

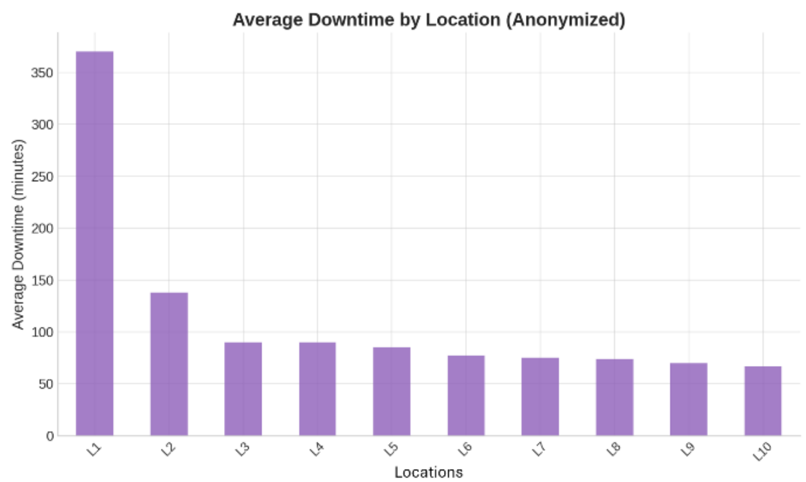
A small set of categories (C1–C3) accounts for the vast majority of events (Figure 9). Concentrating preventive programs and spare-parts inventories on these categories can deliver disproportionate reliability gains.

ISSN: 0718-5073

DOI: https://doi.org/10.7764/RIC.00164.21



(a)



(b)

Figure 8. Downtime Durations and Downtime by Location (Anonymized).

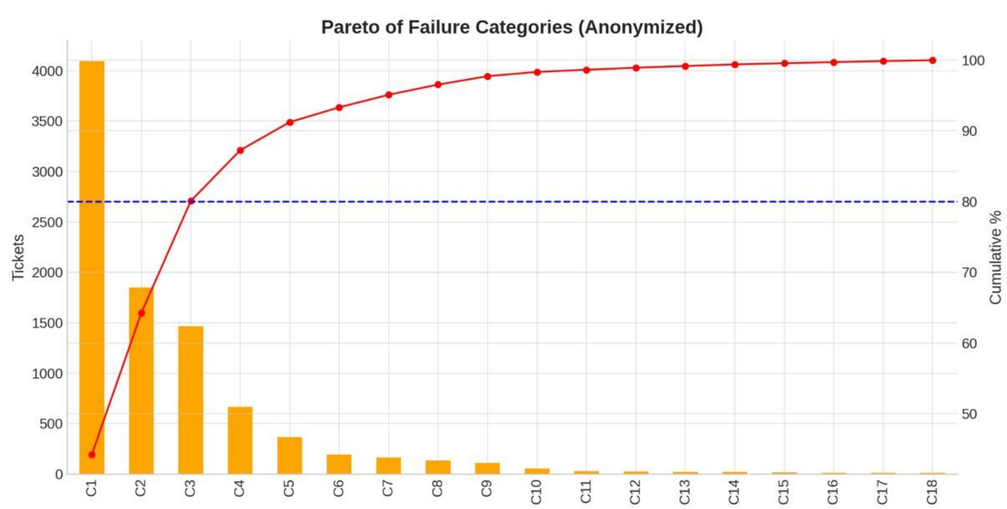


Figure 9. Pareto of Failure Categories (Anonymized).

ISSN: 0718-5073
Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

5. Results and discussion

To assess the impact of the WhatsApp-native AI system, results are analyzed at three levels—technical, organizational, and knowledge-management—revealing both performance gains and the mechanisms behind them.

5.1 Technical Level

Operating the in-group LLM agent on the existing WhatsApp group (~150 participants; >700 messages/week) enabled near-complete capture of maintenance-relevant traffic while preserving the natural flow of conversation in the shared chat (cf. Figures 4–5; Table 1). On average, the system processed 77.5 tickets per day (~542.5 per week; ~3.23 per hour), converting unstructured chat exchanges into structured records in the background.

The effect was immediately visible: acknowledgements, workshop intake, and "ready for redeployment" updates appeared in-thread, restoring a common operational picture that had been diluted by 1-to-1 chatbot form submissions. Instead of manually consolidating over 700 weekly messages, administrators received tickets directly in the project format, with data-entry work reduced by ~90%. A task that previously took 3–5 minutes of manual transcription per ticket now requires less than 30 seconds of review, cutting daily typing effort from roughly 4–6.5 hours to under 40 minutes.

Anchoring outputs to controlled dictionaries (assets, locations, event types) stabilized identifiers, enabling longitudinal KPIs and faster issue resolution. Multimodality proved especially valuable: images provided the most reliable machine cues, and by detecting visible codes or inferring families, the agent minimized follow-ups and improved match rates. In cases of lower confidence—such as slang, overlapping threads, or low-quality images—lightweight guardrails (thresholding, short clarifying prompts, or brief human review) ensured that data quality remained high while the natural flow of chat was preserved.

5.2 Organizational Level

Automation reshaped the administration workflow. Instead of manual consolidation across calls, e-mails, and chat messages, administrators shifted to exception-focused supervision—reviewing the minority of low-confidence cases, correcting outliers, and prioritizing tasks.
The economic effect was likewise substantial: by improving coordination and reducing idle time within repair and maintenance cycles, communication flowed more effectively, and machines returned to service sooner. Downtime costs comprise not only rental/ownership charges during stoppage but also the foregone productivity of the equipment. In aggregate, savings were on the order of five figures in euros.
Adoption was driven foremost by change-management zero: instead of moving people to a new tool, the tool was brought to the channel where coordination already lives—the existing WhatsApp group. Established rhythms and norms were preserved, so participation costs stayed negligible while the record became computable in the background.
A second factor was transparency in the native channel. Because acknowledgements and status changes appeared in-thread, a shared operational picture emerged without additional rituals or systems. This visibility tightened handoffs, reduced idle time, and created auditability as a by-product of everyday work rather than a separate task.

5.3 Knowledge Management Level

The use of LLMs to structure data directly at the source proved transformative. Humans are, in effect, LLM-natives: they communicate in free text, photos, and voice, by meeting information at that level and enforcing a strict schema, the system converted raw conversation into analysis-ready data without templates or rigid forms.

This represents a practical shift in data engineering: high-quality operational data was captured with minimal ceremony, enabling reliable KPIs, downtime analytics, and process monitoring—capabilities historically difficult to sustain in environments dominated by unstructured messaging.

From a governance standpoint, the design maintained a positive data-protection posture: processing occurred only inside closed groups, actions were visible in-thread, and records were stored/exported in a GDPR-compliant manner with full traceability.

Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

5.4 Lean Construction and Data-Driven Management

Viewed through a theoretical lens, the system's impact aligns closely with principles of Lean Construction and data-driven management. The automation of transcription and ticket generation directly supports Lean's goal of eliminating waste, specifically by reducing non-value-adding work and freeing up human capacity for problem-solving and process improvement. The resulting reduction of idle time and faster turnaround cycles demonstrates Lean's emphasis on flow efficiency and continuous improvement in real operational settings.

Simultaneously, the system operationalizes data-driven management: informal exchanges become structured inputs supporting transparency, prediction, and continuous improvement. This creates a direct pathway from everyday communication to measurable productivity gains.

6. Conclusions

This study shows that the central barrier to digitalization in construction is not technological capability but friction in everyday workflows. Embedding an LLM-based agent directly in the existing WhatsApp group, rather than moving crews to external forms or separate tools, eliminated that friction. Under real operating conditions, the system achieved near-complete capture of maintenance-relevant communication, reduced manual transcription by ~90%, restored shared situational awareness in-thread, and contributed to shorter maintenance cycles.

The results illustrate a practical shift in data engineering for field operations: a semantics-first pipeline that meets workers at the level where meaning is naturally produced and binds the output to a strict schema with light human oversight. This enabled reliable, analysis-ready data (Figures 6–9) without additional apps, logins, or behavioral change, while maintaining traceability and data-protection safeguards.

Two mechanisms drove productivity effects:
1. Automating routine data capture within the communication channel already used for coordination,
2. Strengthening decision-making by generating continuous, structured information that reveals operational rhythms, bottlenecks, and high-impact failure categories.

These findings demonstrate that AI can accelerate digitalization and data-driven management when deployed inside the social and operational infrastructure that workers already rely on.

The broader lesson is architectural: systems scale when technology adapts to people, not the reverse. Low-friction, in-channel AI turns everyday conversation into computable data, enabling productivity gains that are both measurable and achievable under real project constraints.

7. Declaration of generative AI and AI-assisted technologies

In preparing this manuscript, AI-assisted technologies were used to support the structuring of the text and improvement of clarity, readability, and presentation. All outputs generated with the assistance of these technologies were carefully reviewed and edited by the authors, who take full responsibility for the content of the manuscript.

8. Notes on Contributors

Diego Cisterna, Co-Founder and Managing Director, Valoon GmbH, Karlsruhe, Germany. **ORCID** https://orcid.org/0000-0003-4282-1141	**Marc Weinmann**, Research Associate, Karlsruhe Institute of Technology, Karlsruhe, Germany. **ORCID** https://orcid.org/0000-0003-3871-0553
Sebastián Cruz, Co-Founder of Crucerlabs, Developer, Queretaro, Mexico.	

9. References

Agarwal, R.; Chandrasekaran, S.; Sridhar, M. (2016). Imagining Construction's Digital Future. McKinsey & Company. https://www.mckinsey.com/~/media/mckinsey/business%20functions/operations/our%20insights/imagining%20constructions%20digital%20future/imagining-constructions-digital-future.pdf

Vol. 40
Special Issue 2025
DOI: https://doi.org/10.7764/RIC.00164.21

Álvarez, T. P.; Barkokebas, B.; Ibáñez, A. J. P.; Costa, D. B. (2024). BIM and procurement data integration in industrialized construction using artificial intelligence. Revista Ingeniería de Construcción, 39(3). https://doi.org/10.7764/RIC.00113.21

BauInfoConsult. (2020, July 8). Social Media am Bau: Stärkerer Einsatz von WhatsApp im Berufsalltag. https://bauinfoconsult.de/social-media-am-bau-staerkerer-einsatz-von-whatsapp-im-berufsalltag/

Borowiecki, M.; Pareliussen, J.; Glocker, D.; Kim, E. J.; Polder, M.; Rud, I. (2021). The impact of digitalisation on productivity: Firm-level evidence from the Netherlands (OECD Economics Department Working Papers No. 1680). Organisation for Economic Co-operation and Development. https://one.oecd.org/document/ECO/WKP(2021)31/En/pdf

Cisterna, D.; Lauble, S.; Haghsheno, S.; Wolber, J. (2022). Synergies Between Lean Construction and Artificial Intelligence: AI Driven Continuous Improvement Process (p. 210). https://doi.org/10.24928/2022/0122

Deloitte. (2025). Deloitte's Global Powers of Construction report finds that the rebounding construction industry expects significant growth through 2030. https://www.deloitte.com/global/en/about/press-room/deloitte-global-powers-of-construction-report.html

Deubel, M. (2021). Untersuchungen zur Wirtschaftlichkeit von Building Information Modeling (BIM) in der Planungs- und Realisierungsphase von Bauprojekten. Karlsruher Institut für Technologie (KIT).

Khan, A. A.; Bello, A. O.; Arqam, M.; Ullah, F. (2024). Integrating Building Information Modelling and Artificial Intelligence in Construction Projects: A Review of Challenges and Mitigation Strategies. Technologies, 12(10), 185. https://doi.org/10.3390/technologies12100185

Kyivska, K.; Tsiutsiura, S. (2021). Implementation of artificial intelligence in the construction industry and analysis of existing technologies. Technology Audit and Production Reserves, 2, 12–15. https://doi.org/10.15587/2706-5448.2021.229532

McKinsey & Company. (2017). Reinventing Construction: A Route to Higher Productivity. https://www.mckinsey.com/~/media/McKinsey/Industries/Capital%20Projects%20and%20Infrastructure/Our%20Insights/Reinventing%20construction%20through%20a%20productivity%20revolution/MGI-Reinventing-Construction-Executive-summary.ashx

Mohd Nawi, M. N.; Baluch, N.; Bahauddin, A. Y. (2014). Impact of Fragmentation Issue in Construction Industry: An Overview. MATEC Web of Conferences, 15, 01009. https://doi.org/10.1051/matecconf/20141501009

Papadonikolaki, E.; Krystallis, I.; Morgan, B. (2020). Digital transformation in construction-Systematic literature review of evolving concepts.

Puolitaival, T.; Kestle, L.; Kähkönen, K. (2018). What's the real story around digital technologies in construction management?

Regona, M.; Yigitcanlar, T.; Xia, B.; Li, R. Y. M.; Regona, M.; Yigitcanlar, T.; Xia, B.; Li, R. Y. M. (2022). Opportunities and Adoption Challenges of AI in the Construction Industry: A PRISMA Review. Journal of Open Innovation: Technology, Market, and Complexity, 8(1). https://doi.org/10.3390/joitmc8010045

Stojanovska-Georgievska, L.; Sandeva, I.; Krleski, A.; Spasevska, H.; Ginovska, M.; Panchevski, I.; Ivanov, R.; Perez Arnal, I.; Cerovsek, T.; Funtik, T. (2022). BIM in the Center of Digital Transformation of the Construction Sector—The Status of BIM Adoption in North Macedonia. Buildings, 12(2), 218. https://doi.org/10.3390/buildings12020218

Tian, D.; Li, M.; Shi, J.; Shen, Y.; Han, S. (2021). On-site text classification and knowledge mining for large-scale projects construction by integrated intelligent approach. Advanced Engineering Informatics, 49, 101355. https://doi.org/10.1016/j.aei.2021.101355

Wu, C.; Li, X.; Guo, Y.; Wang, J.; Ren, Z.; Wang, M.; Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. Automation in Construction, 134, 104059. https://doi.org/10.1016/j.autcon.2021.104059

Yahoo Finance. (2025, July 16). Germany Construction Industry Report 2025 | Output to Register an AAGR of 3.1% During 2026-2029, Supported by the Government's Investments in Transport and Energy Projects. https://finance.yahoo.com/news/germany-construction-industry-report-2025-100900569.html

Zhang, F.; Chan, A. P. C.; Darko, A.; Chen, Z.; Li, D. (2022). Integrated applications of building information modeling and artificial intelligence techniques in the AEC/FM industry. Automation in Construction, 139, 104289. https://doi.org/10.1016/j.autcon.2022.104289