Driving Process Intelligence in Construction Project Operations through Process Mining

Araham Jesus Martinez Lagunas

A Thesis

In the Department

of

Building, Civil, and Environmental Engineering (BCEE)

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy in Building Engineering at

Concordia University

Montreal, Quebec, Canada

March 2026

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Araham Jesus Martínez Lagunas

Entitled: **Driving Process Intelligence in Construction Project Operations through Process Mining**

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN BUILDING ENGINEERING

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

____Chair
Dr. Rabin Raut

____External Examiner
Dr. Daniel Amyot

____Examiner
Dr. Amin Hammad *External to Program*

____Examiner
Dr. Fuzhan Nasiri

____Examiner
Dr. Rebecca Dziedzic

____Thesis Supervisor
Dr. Mazdak Nik-Bakht

Approved by ____
Dr. Po-Han Chen - Graduate Program Director

27th March 2026 ____
Mourad Debbabi
Dean, Gina Cody School of Engineering and Computer Science

ABSTRACT

Driving Process Intelligence in Construction Project Operations through Process Mining

Araham Jesus Martinez Lagunas, Ph.D.

Concordia University, 2026

Construction is one of the world's largest economic sectors and directly influences national productivity and societal well-being. Despite its scale and importance, industry productivity has stagnated for decades, with growth rates three to five times lower than those of other sectors, while nearly 75% of construction projects worldwide experience schedule delays and cost overruns. A major contributor to this underperformance is the lack of systematic, data-driven methods that effectively integrate and operationalize digital technologies to support operational performance improvement. Many Architecture, Engineering, Construction, and Facility Management (AEC/FM) organizations remain predominantly project-oriented and often underestimate the role of systematic process performance monitoring and management in achieving consistent project delivery. As a result, the industry continues to rely largely on manual and subjective methods (e.g., workshops, interviews, focus groups) to capture and manage construction processes, which fail to reflect their actual "as-happened" execution, dynamic evolution, and operational performance. In the absence of such insights, process improvement decisions remain intuitive, judgment-based, and disconnected from empirical evidence. Moreover, AEC/FM organizations increasingly seek to harness project operations data from siloed source systems to enable automated extraction of dynamic process performance insights; however, most Common Data Environments (e.g., Procore) remain largely static and lack such dynamic process monitoring capabilities. Enabling this capability requires the integration and transformation of raw data-oriented system structures into well-defined, process-aware structures.

To address these challenges, this research proposes an automated, data-driven method for quantitatively assessing the performance of inter-organizational construction business processes by integrating process mining techniques with Lean-based quantitative metrics, including cycle time, work in progress (WIP), and lead time. This hybrid Lean-based Process Mining and Management (LPMM) methodology comprises five levels and seven interrelated, process-oriented

modules that follow a stepwise progression, transforming data-oriented architectural layers into process-aware structures. Module 1 establishes prerequisites through automated data extraction, integration, and preparation from heterogeneous project information systems, supported by exploratory data analysis, task mining, and expert consultation. Module 2 translates research objectives into analytical questions by defining performance dimensions, key Process Performance Indicators (PPIs), abstraction levels, and stakeholder perspectives, informed by systematic literature review and domain expertise. Module 3 enables automated event log construction by transforming integrated data into standardized, process-aware representations through automated extractors, schema transformers (e.g., XES), loggers, and enrichment mechanisms. Building on this foundation, Module 4 applies process mining techniques, including Fuzzy Miner and Inductive Miner, to automatically discover "as-happened" end-to-end (E2E) process models. Module 5 derives process efficiency formulations grounded in process mining theory and applies quantitative performance metrics across real-world construction projects to evaluate efficiency, detect deviations, analyze process variants, and assess the impact of Request for Information (RFI) content using natural language processing (NLP). Finally, Modules 6 and 7 translate findings into actionable recommendations and enable continuous monitoring through custom process boards.

The framework was applied to the RFI process across 71 real-world projects comprising 5,564 RFIs, including a cross-case analysis of two large commercial projects. The application involved automated data extraction from the Common Data Environment (CDE), integration and transformation into enriched event logs, portfolio-level efficiency assessment, and detailed process-level analysis of time, cost, control-flow, and team dynamics. Results show that process efficiency among similar projects executed by the same general contractor can vary substantially due to RFI flow length, team dynamics, and staff availability. Improved flow control could have tripled efficiency gains and generated savings of approximately $114,000 (USD) on a single commercial project with an estimated value of $50 million (USD). Scaled across 71 projects, the findings reveal significant opportunities for cost reduction, productivity improvement, and evidence-based decision-making. In total, 42 targeted process improvement recommendations were derived. The extracted as-happened model achieved a fitness of 0.958, a precision of 0.761, and an F-score of 0.845, confirming its representational suitability. Overall, LPMM enables AEC/FM organizations to transform project operations data into process knowledge assets for automated process health monitoring and continuous performance improvement.

ACKNOWLEDGEMENTS

Reaching this professional milestone would not have been possible without the shelter of Your protection, grace, strength, and love. Thank You with all my heart, my beloved God, especially for giving me the strength to fight and overcome COVID-19 back in 2020 when no vaccines existed, and to be a support for my family during those stressful and difficult times. All honor, glory, and victory are Yours!

I would also like to express my deepest gratitude to my PhD supervisor and mentor, Professor Mazdak Nik-Bakht. Thank you for boosting my professional career since day one of the program, and even before, during the interview selection period, and for trusting in my potential with your continuous support all the way. Our shared passion for construction, including this research topic, along with your constructive critiques, high standards, leadership, and continuous feedback, have been essential in shaping this work. Throughout this journey, I have not only learned advanced technical knowledge and practical skills from you, but also critical thinking and communication skills. Whenever things seemed uncertain or during challenging moments, you remained patient in guiding and supporting my efforts. For that and so much more, my sincere thanks and deepest respect. You continuously inspire me to excel and grow in my professional career.

I would also like to acknowledge all the for the support received from all the members of my evaluation committee, Professor Amin Hammad, Professor Rebecca Dziedzic, and Professor Fuzhan Nasiri. Profound thanks for all your thorough reviews, constructive feedback, and precious advice across all the different phases of this program, which have definitely guided and contributed to the achievement of the defined research goals with added value. My sincere thanks are extended to my external examiner, Professor Daniel Amyot, for his commitment to serving in this role and for his detailed evaluation of my thesis. I greatly value his professionalism, integrity, academic insight, and thoughtful feedback, which enriched this work. My sincere thanks also go to Jenny Drapeau and Elena Castiglione, my Graduate Program Advisor and Office Assistant, respectively, for their continued support.

I also acknowledge and deeply thank for the financial support provided from Concordia University and from Professor. Mazdak Nik-Bakht research funds during the first years of the PhD program.

The practical application component of this research would not have been possible without the involvement and support of our industry partner, Aedo AI, to whom I express my sincere thanks, especially to the Aedo team members with whom I collaborated during the three NSERC-Mitacs internships: Chris, Liam, Taylor, Farzaneh, Soroush Salehi, Sarvin, Mark, Ryan, Curtis, Hansen, Sheila, and Shannon together with the Compleccity Lab research members involved in the Aedo project: Daria, Nariman, Soroush, and Amir. I also thank the Mitacs Concordia interns, Vibhav, Gurjote, and Adrien, for their collaboration in the project.

I also cannot thank enough my colleagues from the Compleccity team. Special thanks go to Arash Hosseini, Rafaela Paniza, Julia Gomes, Soroush Abbaspour, Soheil Heidarian, and Azadeh Soofypour, as well as to all my colleagues from the Construction Automation Lab, particularly Abdelhady Omar, Hassan Bardareh, and Dena Shamsollahi for their support and encouragement. I value the journey we shared together!

During the final phase of my PhD studies, Pomerleau supported my professional development and provided the final push to complete my program. I am truly grateful to Pomerleau for the flexibility and ongoing support that made it possible for me to work part-time while completing the final requirements of my PhD, an investment in human potential and long-term vision that only a few outstanding companies embrace. Special thanks go to Eric Lessard, Tarek Gouda, Fernando Valdivieso, Dragica Adamovic, Mathieu Asselin, and Jeremy Mols.

I am also deeply grateful to the members of the Canadian Process Mining Community, especially Professors Wil van der Aalst, William Wan Woensel, Daniel Amyot, Nick Blackbourn, Soren Frederiksen, Marco Montali, and Shahin Karimidorabati, with whom I had the privilege to collaborate and learn from. I also thank process mining software vendors such as Celonis, Mindzie, Disco, Apromore, and others for providing academic licenses that enabled advanced data-driven process analytics.

My heartfelt thanks also go to my beloved family: my beautiful mother, Toñita Lagunas, whose life brightens and guides mine, and my dear father, Miguel Angel Martinez, whose strong mindset and advice I deeply admire, as well as my siblings Kristel, Vanessa, Miguel, Elisabeth Reyes (my sister at heart). Your continuous encouragement and prayers have been the foundation that helped me stay resilient and navigate this challenging and exciting journey. My deepest gratitude goes to my wonderful and lovely sister Kris, who has encouraged me every step of the way and generously dedicated her time to providing advice and support in better shaping and communicating my research. Also big thanks to my beautiful aunts Hildita, Chelita, and Carmelita Lagunas for their prayers and continued support. I feel truly blessed to have you in my life, my dear family. To my beloved grandparents, Ruben Lagunas, Toñita Ruiz, Adelita Contreras, and Jesus Martinez. Ruben and Adelita witnessed the beginning of this journey before they passed away, and I carry their love and memory in this achievement. This milestone is as much theirs as it is mine.

Last but definitely not least, my profound thanks go to my dearest friends for their constant encouragement and support throughout this journey, especially Jorge Lozano, Stacey Conrad, Marla Caserta, Miguel Ricalde, Julia Gomes, Alex ML, Julia Kalenski, Casaundra, Delaram, Hassan, Reza, Soroush, Rafaela, Zach, Fany, Tan, Lucie, Moe Askari, Zeineb, Adriana Ambrosio, Tavo, Lau, Javi, Karen, Angie, Rodri, Andrea, Zaira, Carlo, and Mizanoor.

DEDICATION

I dedicate this thesis to:

My beloved God.

To my precious parents, Toñita Lagunas and Miguel Martinez, who are the heroes of my life and exceptional role models, guiding me not only with strong values but also by supporting my education since I was young. I couldn't be luckier than to be your son!

To my siblings Vanessa, Kris, and Miguel, my little nephew Miguelito, and to all my beloved family Martinez and Lagunas.

A very special acknowledgment goes to my beloved mother, Toñita, and my dear aunt, Hildita, for their presence at my PhD oral examination in Montreal. Your support gave me strength, confidence, and peace, and deeply motivated me to perform at my very best throughout the examination.

To those who have inspired me throughout my life by their lifelong dedication to helping, teaching, and serving others, while never forgetting about themselves!

To the readers!

Dedico esta tesis a:

Mi amado Dios.

A mis queridos padres, Toñita Lagunas y Miguel Martinez, quienes son los héroes de mi vida y modelos excepcionales a seguir, guiándome no solo con sólidos valores, sino también apoyando mi educación desde que era pequeño. ¡No podría ser más afortunado de ser su hijo!

A mis amados hermanos Kris, Vanessa, y Miguel, a mi sobrino Miguelito, y a toda mi amada familia Martinez y Lagunas.

¡A quienes me han inspirado a lo largo de mi vida con su forma de vivir para ayudar, enseñar y servir a los demás!

¡A los estimados lectores!

CONTRIBUTION OF AUTHORS

Authorship contribution statement

Araham Jesus Martinez Lagunas: Conceptualization, comprehensive literature review, data extraction, data curation, data integration/fusion, data processing, LPMM methodological framework design, framework implementation, investigation, data and process analytics, algorithm development for event log generation, data-driven process assessment, framework validation, visualizations, writing full original draft, review and editing.

Mazdak Nik-Bakht: research conceptualization, funding acquisition, industry partner engagement, principal project investigator, supervision, writing review.

All authors reviewed the final manuscripts and approved their content for publication.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

Acronym	Definition
AEC/FM	Architecture, Engineering, Construction / Facility Management
AI	Artificial Intelligence
API	Application Programming Interface
BAS	Building Automation System
BIM	Building Information Modeling
BPM	Business Process Management
BPMN	Business Process Management and Notation
BPR	Business Process Reengineering
CCPM	Conceptual Construction Process Model
CDE	Common Data Environment
CDM	Common Data Model
CMIS	Construction Management Information Systems
CMMS	Computerized Maintenance Management Systems
CPS	Cyber-Physical Systems
CSV	Comma Separated Values
D2P	Data-to-Process
DES	Discrete Event Simulation
DFG	Directly-follows Graph
DIKW	Data–Information–Knowledge–Wisdom
DM	Data Model
DMAIC	Define, Measure, Analyze, Improve, Control
e-EPC	Extended Event-driven Process Chain
E2E	End-to-End
EDA	Explanatory Data Analysis
EKG	Event Knowledge Graphs
ERD	Entity Relationship Diagram
ERP	Enterprise Resource Planning
ETL + E	Extract, Transform, Load, and Enrich
FOL	First-Order Logic

FPDMs	Foundational Process Data Models
GC	General Contractor
GDP	Gross Domestic Product
GPT	Generative Pre-Trained Transformers
HMI	Human–Machine Interactions
HTTP	HyperText Transfer Protocol
ICT/IT	Information and Communication Technology
IDEF	Integration Definition
IFC	Industry Foundation Classes
IFP	Industry Foundation Processes
IM	Inductive Miner
IoT	Internet of Things
IS	Information Systems
JSON	JavaScript Object Notation
KM	Knowledge Management
KPIs	Key Performance Indicators
LLM	Large Language Models
LPG	Labeled Property Graph
LPMM	Lean-based Process Mining and Management
MEP	Mechanical, Electrical, and Plumbing
NLP	Natural Language Processing
OCDM	Object Centric Data Model
OCED	Object Centric Event Data
OCEL	Object Centric Event Logs
OCPM	Object Centric Process Model
OR	Owner Representative
OWL	Web Ontology Language
PAIS	Process-aware Information Systems
PC	Project Coordinator
PDCA	Plan–Do–Check–Act
PHM	Process Health Monitoring
PM	Project Manager
PMIS	Project Management Information Systems
PMM	Process Modeling and Management
PPIs	Process Performance Indicators

PQL	Process Query Language
PSM	Performance Spectrum Miner
PSO	Particle Swarm Optimization
RAG	Retrieval-Augmented Generation
RDF	Resource Description Framework
REST	Representational State Transfer
RFI	Request For Information
RFID	Radio Frequency Identification
RFQ	Request For Quotation
ROI	Return on Investment
RPA	Robotic Process Automation
SC	Specialty Contractor or Subcontractor
SNA	Social Network Analysis
SOP	Standard Operating Procedures
SQL	Structured Query Language
UAV	Unmanned Aerial Vehicles
VAT	Value Added Time
WF	Workflow Foundation
WfMS	Workflow Management Systems
WIP	Work in Progress
XES	eXtensible Event Stream
XML	eXtensible Markup Language

PREFACE & LIST OF PUBLICATIONS

This thesis is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Building Engineering at Concordia University. The research was conducted under the supervision of Dr. Mazdak Nik-Bakht, Associate Professor in the Department of Building, Civil, and Environmental Engineering (BCEE).

This dissertation consolidates the research carried out during the PhD program, documented in the following technical publications:

Peer-reviewed Journal Papers

1. A. J. Martinez Lagunas and M. Nik-Bakht. "Process Mining, Modeling, and Management in Construction: A Critical Review of Three Decades of Research Coupled with a Current Industry Perspective." *Journal of Construction Engineering and Management.* vol. 150, no. 11. 2024. https://doi.org/10.1061/jcemd4.coeng-14727.
2. A. J. Martinez Lagunas and M. Nik-Bakht. 2025. "Data-Driven Insights for Enhancing RFI Process Performance in AEC/FM Projects". *Journal of Construction Engineering and Management*. **Status:** Submitted - Under Review.
3. A. J. Martinez Lagunas and M. Nik-Bakht. 2026. "An Automated Data-to-Process (D2P) transformation method for RFI process performance assessment and monitoring." *Automation in Construction*. **Status:** Submitted - Under Review.
4. S. Abbaspour, A. J. Martinez Lagunas, and M. Nik-Bakht. 2025. "Enhancing Contractor Financial Performance by Leveraging Maintenance Orders' Data." *Journal of Management in Engineering.* **Status:** Submitted 2nd Round - Under Review.

Peer-reviewed Conference Papers

5. Martinez, A., Nik-Bakht, M. (2024). "Enabling Process Mining in the Construction Industry: An Event Log Schema for Change Management Process." In: Desjardins, S., Poitras, G.J., Nik-Bakht, M. (eds) Proceedings of the Canadian Society for Civil Engineering Annual Conference 2023, Volume 3. *CSCE 2023*. Lecture Notes in Civil Engineering, vol 497. Springer, Cham. https://doi.org/10.1007/978-3-031-62170-3_8
6. Martinez Lagunas, A. J. & Nik-Bakht, M. (2024). "Analysis of XES and OCEL data schemas: Towards multidimensional process mining of intertwined construction processes." *Waterloo: IAARC Publications.* Retrieved from: https://www.proquest.com/docview/3092410989
7. Martinez Lagunas, A. J., Abbaspour, S., & Nik-Bakht, M. (2024). "NLP for Automated Discovery and Assessment of Dominant Construction and Maintenance Work Order Activities in MEP Projects." *Construction Research Congress 2024: Advanced*

Technologies, Automation, and Computer Applications in Construction. Des Moines, Iowa. https://ascelibrary.org/doi/10.1061/9780784485262.018

8. Martinez Lagunas, A.J., Askarihosni, M., Alimohammadi, N., Dezyanian, A., Nik-Bakht, M. (2023). "Discovery of Energy Performance Patterns for Residential Buildings Through Machine Learning." In: Walbridge, S., et al. Proceedings of the Canadian Society of Civil Engineering Annual Conference 2021. *CSCE 2021*. Lecture Notes in Civil Engineering, vol 247. Springer, Singapore. https://doi.org/10.1007/978-981-19-0968-9_1

9. Nabipour, N., Martinez, A., Nik-Bakht, M. (2025). "Predicting Change Order Magnitude in Construction Projects: A Machine Learning Approach." In: Desjardins, S., Poitras, G.J., Nik-Bakht, M. (eds) Proceedings of the Canadian Society for Civil Engineering Annual Conference 2023, Volume 4. *CSCE 2023.* Lecture Notes in Civil Engineering, vol 498. Springer, Cham. https://doi.org/10.1007/978-3-031-61499-6_11

10. Abbaspour, S., Martinez, A., Sandhu, G.S., Nik-Bakht, M. (2024). "Automated Identification and Impact Quantification of Financial Budget Items from Construction Data." In: Desjardins, S., Poitras, G.J., Nik-Bakht, M. (eds) Proceedings of the Canadian Society for Civil Engineering Annual Conference 2023, Volume 3. *CSCE 2023.* Lecture Notes in Civil Engineering, vol 497. Springer, Cham. https://doi.org/10.1007/978-3-031-62170-3_1

Industry White Paper

11. A. J. Martinez Lagunas, S. Abbaspour, M. Nik-Bakht, and M. Ouf, "Digitalization and Process Management in Construction: A Current Industry Perspective," CMAA, 2025. https://www.cmaanet.org/sites/default/files/resource/Digitalization%20FINAL.pdf

CHAPTER 1. INTRODUCTION

1.1. Background

Construction is one of the world's largest economic sectors, contributing over USD 10 trillion annually to the global GDP and employing more than 7% of the workforce worldwide [1]. The sector plays a key role in building, maintaining, and retrofitting quality infrastructure, which serves as the backbone of global competitiveness, urban sustainability, and economic growth. This sector directly influences a nation's productivity and its society's well-being.

Despite these aspects, the construction industry has historically underperformed in terms of productivity, growing in the last three decades three to five times slower than other industries such as manufacturing, information technology (IT), wholesale and retail, healthcare, and finance [2-4]. A key driver boosting productivity in these sectors is their ability to advance their digital transformation in an agile manner by effectively identifying, adopting, and implementing novel technologies and methodologies to enhance operational performance [5-9]. The Fourth Industrial Revolution, or Industry 4.0, is marked by several fast-paced developments in the area of Information and Communication Technology (ICT), including, among others, the Internet of Things (IoT), Cyber-Physical Systems (CPS), Digital Twins, Building Information Modeling (BIM), Robotic Process Automation (RPA), Artificial Intelligence (AI), cloud and quantum computing, Business Process Management (BPM), big data analytics, and process mining, [10-13].

These novel technologies and methodologies have brought with them a series of challenges and opportunities for business organizations. On the one hand, they have forced industries to transform their traditional supply chains (i.e., siloed and static) into digitalized supply chains (i.e., integrated and dynamic) with a special focus on operational management and process performance traceability. On the other hand, they have helped companies in improving their productivity levels resulting in a significant advantage over competitors [14]. Similarly, 'Construction 4.0' aims to adopt latest ICT developments in the construction industry to enhance operational efficiency, increase productivity, and support decision-making across the lifecycle of construction projects in a more objective, fact-based, and data-driven format [15-17].

Construction projects are not meant to last indefinitely; they are constrained by time, budget, manpower, quality requirements, and other factors. These projects are dissolved once the agreed-upon scope of work is completed and the established goals are met in accordance with the contractual agreement between the parties involved. Therefore, the execution of these projects must be efficient to ensure on-time project delivery within the planned budget, and in conformance with client requirements and technical specifications. Additionally, for construction firms, projects should be profitable and serve as mechanisms for continuous performance improvement and innovation, while opening up new market opportunities [18, 19]. Yet, despite the extensive experience of some construction organizations in managing ‘projects’, nearly 75% of construction projects worldwide face delays and cost overruns [20, 21], with companies continuously struggling with severe productivity challenges and poor operational performance [19, 22, 23]. One significant factor contributing to this figure is that construction organizations often tend to underestimate the importance of effectively monitoring and managing the performance of their intra- and inter-organizational business ‘processes’ required for successful project execution and delivery. As a result, business operations become inefficient, ultimately leading to loss of project control and declining productivity [24-26].

Organizations in the Architecture, Engineering, Construction, and Facility Management (AEC/FM) domain seeking to enhance project portfolio performance and maintain competitiveness must recognize and address the inherently process-intensive nature of construction projects. Inefficient execution of critical processes poses a significant risk to overall project delivery [27]. Therefore, capturing essential business processes (either operational or administrative) and extracting data-driven performance insights from them in an automated manner are important items on the digital transformation agenda and a prerequisite towards construction automation [26].

1.2. Research Motivation

According to the McKinsey Industry Report in 2017, improving construction productivity could add 1.6 trillion dollars to the global economy [1]. One of the most critical levers to realize this opportunity lies in the efficient execution and management of construction business processes. The adoption and implementation of data-driven monitoring and effective control mechanisms aimed at improving process efficiency can yield substantial productivity gains, benefiting not only individual projects but also the industry at large [2, 22, 28].

In the best-case scenario, some construction organizations have defined their normative Standard Operating Procedures (SOPs) or their reference “as-planned” business process models. These SOPs and process models, if exist, aim to provide visibility into their business operations. More specifically, they serve as guidelines or mandates for construction project stakeholders, prescribing how to perform the required business process-related activities throughout the conception, execution, completion, operation, and maintenance phases of the construction project lifecycle. Nevertheless, various large construction companies, and especially small and medium-sized organizations, have not even formally documented these operational processes.

Considering these aspects, the main motivation of this research is to support construction organizations in improving the performance of their business operations by providing them with the capability to automatically extract these actual business process models through the exploitation of their project operations data, not only to map and document these processes, but more importantly to quantitatively evaluate and improve their performance in a continuous manner. To this end, *process mining*, a novel data-driven approach grounded in a rapidly evolving academic field, with broad industry adoption and a strong market presence, has become a crucial asset driving digital transformation and boosting productivity across industries [29, 30]. This proven technology presents a clear opportunity to advance process intelligence capabilities within business organizations by exploiting real-world process executions captured from operational project data, thereby enabling the identification, analysis, and improvement of actual process performance behavior [31].

1.3. Problem Statement

Even though the construction sector has a rich body of knowledge in process modeling, simulation, and management, it still heavily relies on manual and subjective means and methods (i.e., workshops, interviews, focus groups) for capturing, modeling and managing its construction processes. Although these methods help construction organizations to initially document and structure conceptually their reference business processes, the resulting process model depictions are static and only reflect: either (i) the expected “as-planned” behavior of construction operations on how certain processes should be executed by stakeholders to conform with best practices and standard operating procedures (SOPs); or (ii) the ideal “to-be” behavior on how processes should be optimized and re-engineered to better perform. More specifically, existing methods fail to capture and reflect the ‘as-happened” behavior on how processes are actually executed in the real-world, how they dynamically change over time, and how they perform. In the absence of such details, decision-making to improve such processes has remained intuitive, judgment-based, and disconnected from factful insights.

Unlike other industries’ processes, construction processes are frequently characterized by diverse non-standard practices and ineffective process control with poor transparency and visibility into operational performance. Operational knowledge captured through manual instruments is frequently lost when not formally documented (e.g., sketched on a workshop board and subsequently erased). Consequently, these conceptual reference process models often become obsolete, and in many cases they do not even exist. In addition, existing data architectures of most construction organizations lack process-awareness, which hinders process intelligence capabilities, as they have not fully leveraged recent ICT advancements and data-driven, process-oriented management approaches to effectively harness the large volumes of data generated through their day-to-day project operations [26]. Consequently, they remain unable to systematically manage and improve their business processes to reduce waste and increase productivity. Existing process-oriented approaches are time-consuming, resource-intensive, and error-prone, often leading to poorly managed processes, operational inefficiencies, and, in some cases, construction project failure [22].

1.4. Problem Significance

The long-standing productivity stagnation in the construction industry is strongly influenced by a digital transformation gap and poorly managed construction business processes [26]. The labor productivity of digitally intensive sectors such as manufacturing, finance, and wholesale grew by 22.1% over the past two decades. In contrast, construction is classified among the lagging, non-digitally intensive sectors that continue to face persistent productivity challenges, growing only 6.3% during the same time period [32], resulting in severe project delays and cost overruns [3, 33, 34]. Efficient processes are the backbone of successful project management and delivery by streamlining daily operations, enhancing coordination, and ultimately improving overall project performance. However, in practice, these intra- and inter-organizational construction management processes (either administrative or field-level operational) are frequently overlooked, and the performance monitoring of these processes (if any) is performed by manual means and time-consuming. Thus, if not properly managed and controlled, they can lead to productivity loss or stagnation, financial stress, delays, or unsuccessful project delivery [26].

Construction organizations in the AEC/FM domain rarely capitalize on the operational data stored within their adopted Common Data Environments (CDEs) or Construction Management Information Systems (CMIS) to systematically and automatically analyze their core End-to-End (E2E) business processes [35]. Consequently, they lack the capability to continuously monitor the performance of their operations in a systematic and data-driven manner. This limitation impedes the identification of real-world process inefficiencies and hinders evidence-based decision-making aimed at reengineering and automating core business operations. For instance, a lack of process transparency and limited visibility into the actual performance of the construction Change Order Management Process can lead to significant negative consequences such as project delays due to work disruptions [36], cost overruns [37], legal claims and disputes that often lead to lengthy and costly litigations [38], as well as overall productivity decline [39-41]. In this regard, previous studies report average project cost overruns between 5 to 10% due to change orders, and for some poorly managed projects the cost overrun due to construction change orders far exceeds 10% [42, 43]. Putting this into perspective, Oxford Economics reported an estimated total global construction project value of USD 10.7 trillion in 2020. Consequently, the estimated global cost overruns for that year due to changes ranged between USD 535 billion and USD 1.07 trillion [44]. Therefore, enabling this process-awareness capability by addressing this gap is essential for improving project performance, increasing operational efficiency, and ultimately enhancing overall economic outcomes.

1.5. Goal and Objectives

Long-term Goal – The proposed research aims at driving process intelligence in construction projects by automatically extracting actionable performance insights from construction projects' operations data. This will be achieved by leveraging the latest ICT developments to automatically generate dynamic, process-oriented insights that inform evidence-based decision-making and drive improvements in construction process management.

Research Objectives – The overarching goal of this research is to develop a systematic, process-aware framework, grounded in a data-driven methodology to guide and facilitate the implementation of process mining in the construction industry. This framework aims to equip construction organizations and practitioners with the capability to automatically derive, quantitatively assess, monitor, and manage the actual execution performance of core construction processes leveraging project operations data. The main research objectives (ROs) to achieve this goal are to:

RO1. Identify, from the existing literature: (i) the reported methods, techniques, and algorithms for Process Modeling and Management (PMM) in construction; and (ii) the core construction business processes frequently investigated and documented across the entire life cycle of a construction project.

RO2. Conduct a requirements analysis to investigate the significance of the identified processes and to select the main construction process(es) that will be further analyzed in this research study by: (i) capturing process-related knowledge from subject matter experts and previous studies; and (ii) extracting operational insights from project operations data.

RO3. Develop a process mining use case for the construction Change Order Management Process and its intertwined Requests for Information (RFI), considering traditional and multidimensional event log data schemas (i.e., XES and OCEL, respectively).

RO4. Develop a process-aware methodological framework with a system-agnostic nature (in terms of software and data standards) and an underlying data-driven methodology tailored to the AEC/FM domain to enhance process intelligence in construction projects by: (i) automating the extraction of "as-happened" E2E process models; (ii) facilitating dynamic process performance assessment, diagnosis and monitoring; (iii) supporting data-driven decision-making and action-oriented process improvements in an objective manner.

RO5. Apply the proposed framework in real-world case studies implementing the developed process mining use cases for RFI and Change Order Management Processes to monitor and quantify their performance behavior, thereby realizing process performance improvement opportunities and strategies.

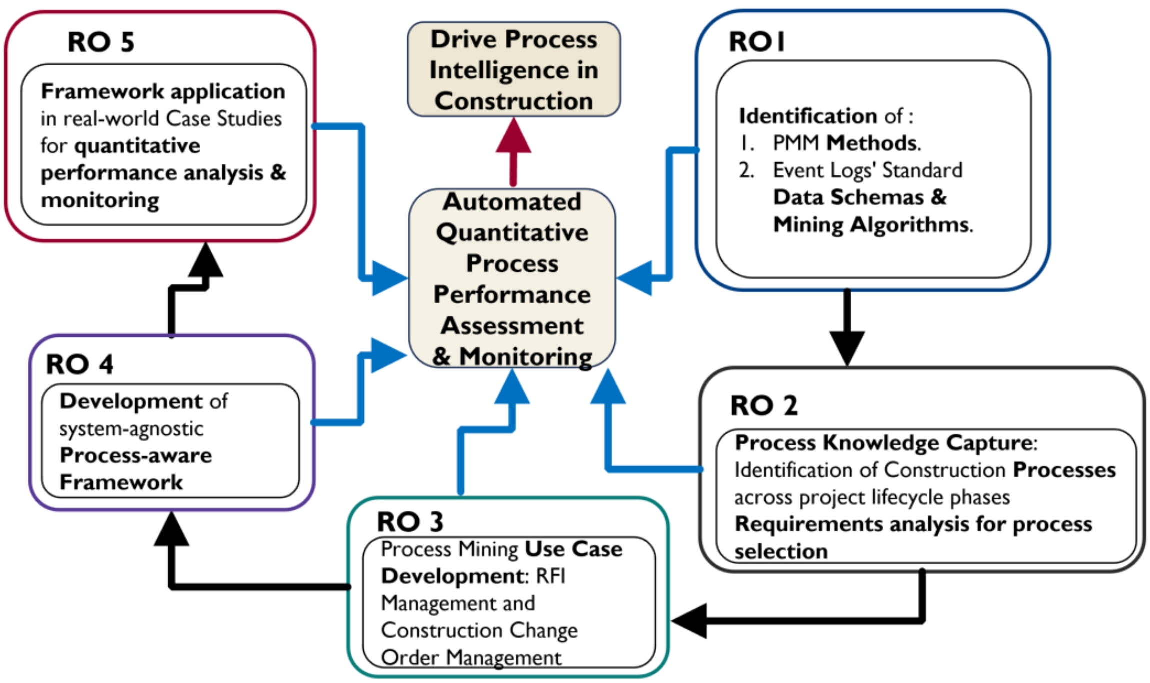


Figure 1-1. Research goal and objectives

By achieving the research objectives previously described and visually outlined in Figure 1-1, this research will primarily contribute to: (i) helping construction organizations harness their project operations systems' data by transforming it into corporate process knowledge assets, thereby shifting from traditional data-aware structures to process-aware architectures; and (ii) providing increased process transparency and visibility with actionable data-driven insights into "as-happened" process performance behavior of real-world core construction business operations.

The expected outcomes of this research will support project managers and directors, project controls specialists, and project coordinators in better controlling operational performance, increasing process efficiency, enhancing overall productivity, and making more informed decisions based on the automated identification, evaluation, and diagnosis of process inefficiencies, which, in turn, can serve as a data-driven foundation for triggering action-oriented process improvement strategies.

1.6. Scope

Grounded in a comprehensive study on process mining, modeling, and management in the AEC/FM domain, this research focuses on the automated extraction and quantitative performance assessment of construction management processes that are recurrently executed across construction projects [26]. These processes are the backbone of business operations within construction organizations; thus, to ensure successful project delivery, it is of paramount importance to effectively control and manage these processes by dynamically assessing their actual execution performance. Such continuous evaluation enables the detection of major operational inefficiencies and the identification of targeted opportunities for process reengineering and automation.

More specifically, this study introduces a novel data-driven methodology that leverages and operationalizes process mining as a capability-enabling mechanism for construction organizations. The approach automates the extraction of E2E "as-happened" inter-organizational construction management processes and enables the dynamic assessment and continuous monitoring of their operational performance. The proposed methodology aims to be generally applicable regardless of the selected construction management process to be assessed and monitored. However, the cross-case study presented in this document focuses on inter-organizational construction management processes that are central to supporting essential project controls functions, including information management, change management, and performance monitoring. These foundational processes are frequently undervalued, even though deficiencies in their execution or management can significantly disrupt project delivery and result in major time and cost overruns [26].

Recognizing these aspects, this research primarily examines the Request for Information (RFI) management process as an upstream process that can initiate and influence downstream project control processes. In particular, RFIs may trigger the construction Change Order Management process, a core business process within project contract administration, as well as related downstream activities such as the execution of construction works associated with approved change orders and the processing of progress payment requisitions. Accordingly, the data-driven analysis presented in this study explicitly focuses on the RFI management process, with downstream interrelated processes considered as consequential outcomes rather than the primary unit of analysis.

Building on this perspective, the effective management of the RFI management process is critical to successful project delivery, as RFIs represent one of the primary upstream mechanisms through which construction changes are initiated. In practice, RFIs frequently trigger the Change Order Management process, a formally defined, contract-based subprocess within project contract administration, particularly during the construction phase. Owing to its formal and contractual nature, the Change Order Management process has significant potential to disrupt construction activities if not properly managed, leading to cost and schedule overruns, reduced stakeholder trust, and, in some cases, contractual disputes and claims. Although largely administered at the contractual or administrative level, change-related processes can substantially affect field operations by hindering work execution, reducing construction productivity, and ultimately degrading overall project performance. For these reasons, this research narrows its scope to the RFI management process as the central subject of analysis, given its pivotal role as an upstream trigger of construction change orders [35, 45-48]. Moreover, this study adopts a diagnostic process analytics perspective rather than a predictive one, aiming to identify, quantitatively diagnose, and monitor major performance inefficiencies in real-world RFI process executions.

1.7. Research Methodology Overview

The high-level research methodology to address the stated problem and to achieve the underlying research objectives is shown in Figure 1-2. It comprised three major components: (i) PMM foundations, featuring a comprehensive literature review on process mining, modeling, and management in the AEC/FM industry that informed the identification of key research gaps and the delineation of this study's scope; (ii) the development of a process-aware framework tailored to the construction domain, including the introduction of a new integrated, data-driven method for the quantitative assessment and dynamic monitoring of construction process performance, which comprises seven process-oriented modules; and (iii) the application of the proposed framework in real-world construction projects through case studies.

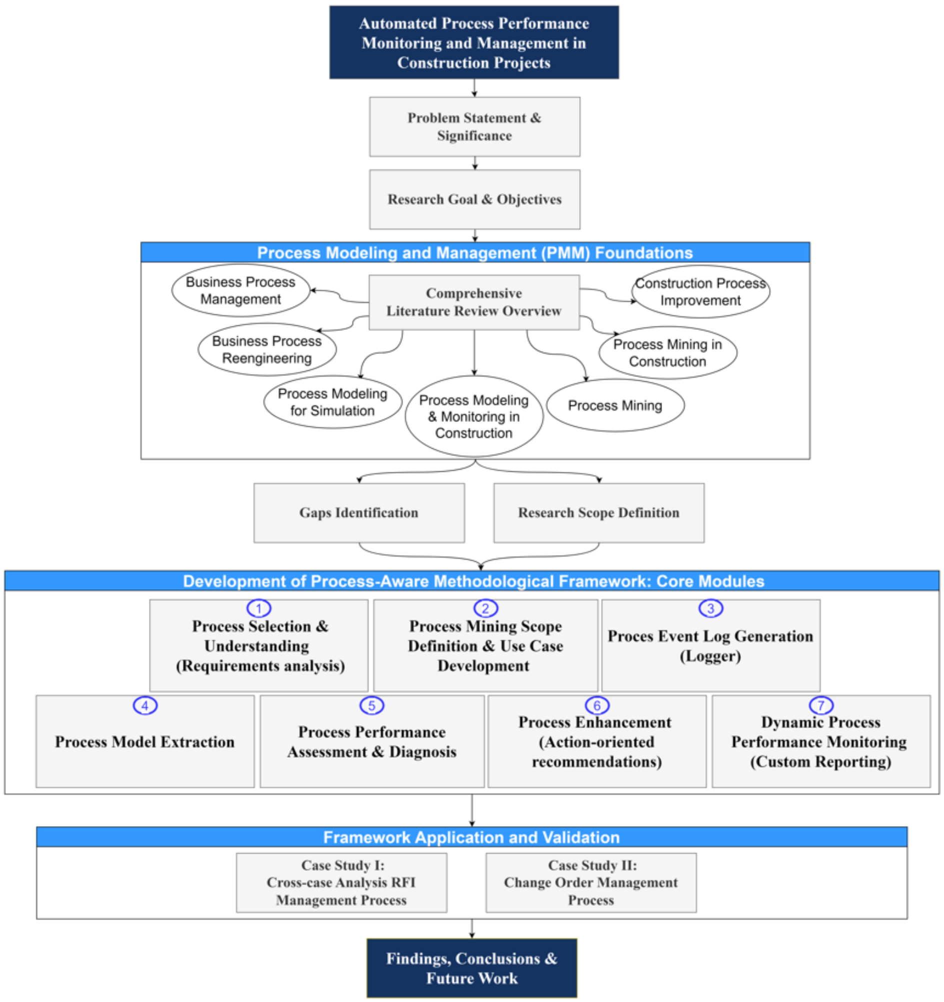


Figure 1-2. High-level research methodology

This study seeks to answer two main research questions:

(RQ1) How can construction organizations harness project operations data to extend and transform their existing data-oriented architectures into process-aware structures to control the performance of their business operations?

(RQ2) How can construction professionals leverage process mining, in combination with other process-oriented methodologies, to extract, quantitatively assess, monitor, and improve the actual E2E performance of critical construction processes in a dynamic, automated, and data-driven manner?

By answering these questions, this research aims to support project directors, managers, coordinators, and project controls specialists in automatically extracting fact-based, actionable insights into operational performance, thereby enabling data-driven decision-making. These insights not only enhance visibility into actual process execution, they also increase transparency by enabling data-driven lifecycle traceability of the parties accountable for executing these critical processes. In this vein, this study embraces Taiichi Ohno's Lean principle, "Look straight at the reality" [49], as this research generally aims to identify and assess actual inefficiencies in the performance of business operations in construction projects through data-driven analysis, with the goal of uncovering actionable and targeted process improvements and automations.

1.8. Thesis Organization

This dissertation comprises eight core chapters, each dedicated to a key aspect of the research. Aside from the introductory chapter (Chapter 1) and the concluding chapter (Chapter 7), the structure of the remaining chapters is presented in Figure 1-3. This figure presents each chapter's objectives, associated methods and techniques, and resulting outcomes, highlighting how each chapter supports the overarching research objectives, as follows:

- Chapter 2 provides an overview of the fundamental concepts underlying PMM by examining existing methods and techniques. Through a systematic review and scientometric analysis, it identifies current PMM practices, understudied research areas within the field, and key theoretical foundations in process mining.
- Chapter 3 focuses on (i) identifying key research gaps and future research directions in PMM, (ii) investigating the most significant intra- and inter-organizational construction business processes across the entire lifecycle of construction projects frequently documented in the literature, and (iii) designing a support tool for documenting, monitoring, and managing these critical processes. Chapters 2 and 3 lay the foundations of this research.
- Chapter 4 centers on the investigation of critical construction processes by analyzing their significance and identifying key operational pain points reported by subject matter experts through semi-structured interviews. It also presents the selection of the most critical process for further analysis, based on a combination of literature review outcomes, expert

input, and data-driven process significance. The latter is determined by analyzing the operational frequency of stakeholder actions related to the execution and management of these processes, captured through task mining analysis in the targeted construction CDE.

- Chapter 5 outlines the design and development of a process-aware framework tailored to the construction domain to guide and support the implementation of process mining in the AEC/FM sector. The framework design is informed by the Data-Information-Knowledge-Wisdom (DIKW) hierarchy [50]. This chapter introduces a newly integrated, data-driven method that combines process mining with Lean-based principles and metrics, along with custom-developed algorithms for the automated generation of process event data logs.

- Chapter 6 delves into the automated extraction and quantitative performance assessment of the actual E2E execution behavior of selected processes, aiming to identify key operational inefficiencies through data-driven analysis. In this regard, this chapter details the application of the developed process-aware framework to real-world construction projects through case study analyses, employing process mining techniques such as the Fuzzy Miner and Inductive Miner algorithms, along with other data-driven methods, techniques, and performance metrics outlined in Figure 1-3.

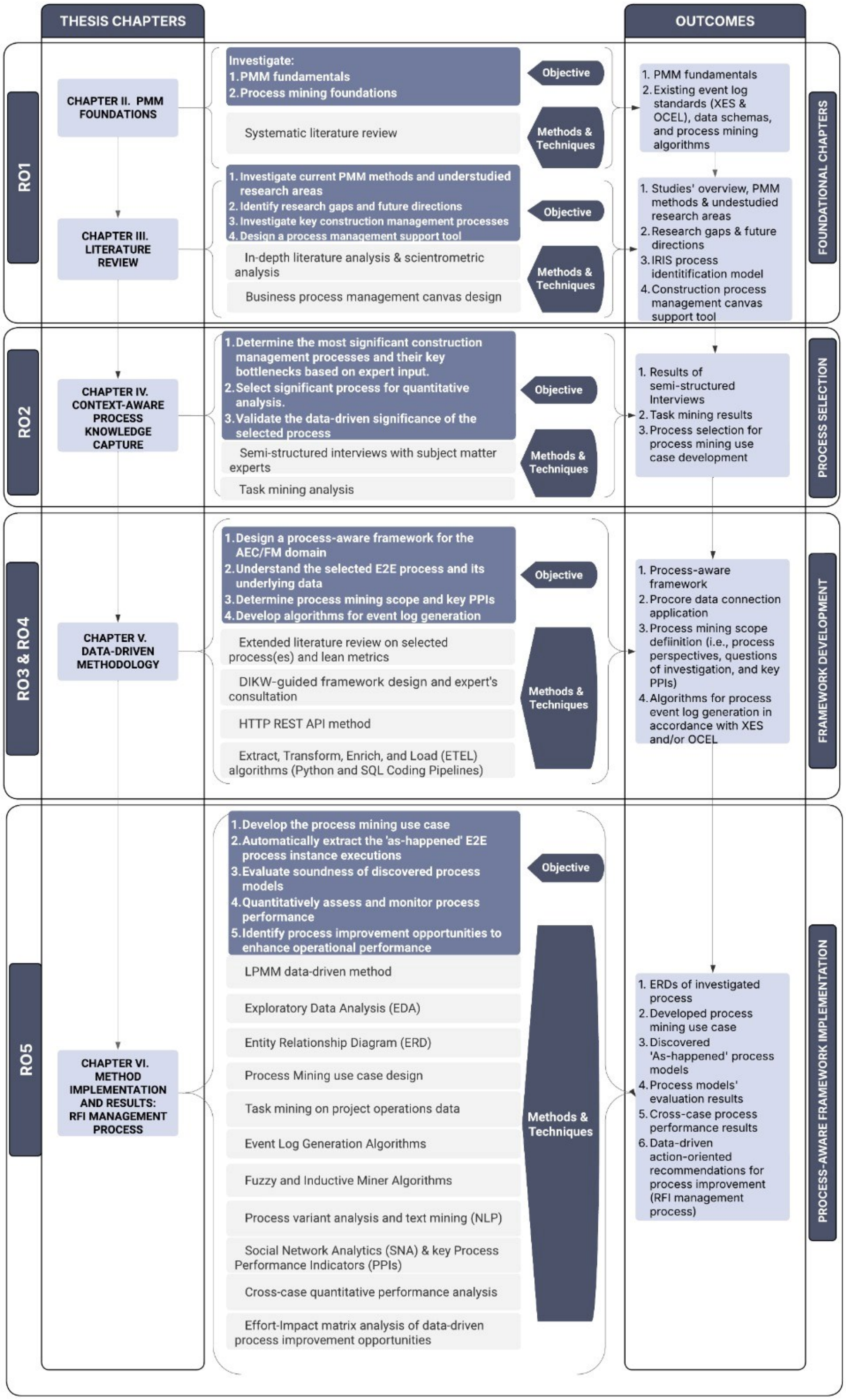


Figure 1-3. Thesis structure

CHAPTER 2. PROCESS MODELING AND MANAGEMENT FOUNDATIONS[1]

2.1. BPM Fundamentals

The adoption of BPM, process mining, and other industry-proven ICT developments, combined with strong process knowledge management (KM) capabilities, is essential for the effective management of business operations [11, 51]. Leveraging novel technologies to capture actual business process executions and extract data-driven performance analytics from them are important items on the digital transformation agenda of the construction industry and a prerequisite for construction automation.

At its most basic definition, a process is referred to as a series of actions taken with a particular goal using resources to transform inputs into outputs. Within an organizational context, a business process is defined as a set of interrelated events and activities whose execution follows a defined sequence. It often involves various objects and multiple actors or stakeholders from one or more organizations, requiring frequent information exchange and multiple key decision points to deliver a valuable product or service to a customer or to achieve an organizational goal [52-54].

BPM falls under the umbrella of “process science” [11], a broad discipline that combines ICT with KM practices to ensure that process-related knowledge (e.g., best practices, lessons learned, and business decision rules) is captured, made accessible, and reused effectively to support and enhance process management. BPM is a management discipline that provides a set of principles, tools, techniques, and comprehensive methodologies focusing on value delivery through process automation and efficient process management [55, 56]. It facilitates process

---

[1] This chapter is based upon:

Martinez Lagunas, A. J., & Nik-Bakht, M. (2024a). Analysis of XES and OCEL Data Schemas: Towards Multidimensional Process Mining of Intertwined Construction Processes. *Proceedings of the 41st International Symposium on Automation and Robotics in Construction*, *Construction Management Techniques*(2413-5844), 444-451. https://doi.org/10.22260/ISARC2024/0058

Martinez Lagunas, A. J., & Nik-Bakht, M. (2024c). Process Mining, Modeling, and Management in Construction: A Critical Review of Three Decades of Research Coupled with a Current Industry Perspective. *Journal of Construction Engineering and Management*, *150*(11). https://doi.org/10.1061/jcemd4.coeng-14727 analysis by supporting, for example, the timely identification of major sources of waste [57, 58]. Moreover, it enables quality assurance capabilities through process modeling with formal semantics, process auditing [59], and process conformance checking the process performance against applicable standards [53, 60, 61]. These BPM capabilities are the lifeblood of organizations, as they strengthen resilience in a rapidly changing industry market landscape, promote operational efficiency, and enrich organizational knowledge memory related to their core business operations [62].

The BPM lifecycle has its roots in the Deming cycle plan-do-check-act (PDCA) model focused on process improvement [63]. The BPM cycle broadens the PDCA model by incorporating advanced stages such as process discovery, modeling, automation, monitoring, and continuous process redesign, extending the focus beyond simple subjective iterative improvements to a more holistic, data-driven extraction, monitoring and management of E2E business processes [64]. The BPM evolution happened across the so-called "three major waves of BPM" [54]. The first wave began in 1911 with F. Taylor and H. Gantt setting the principles of scientific management [65]. The second wave happened in 1990 with the emergence of Business Process Re-engineering (BPR), as an essential process-oriented philosophy to maximize organizational performance by rethinking and redesigning core business operations, which had its boom by the early 90s [66], especially with the study carried by Hammer named "Don't Automate, Obliterate" [67]. Moreover, "The Reengineering Rationale" by Myers (1995) emphasizes the importance of BPR in the construction sector to make informed operational improvements and business-driven strategies enabled by technology [68]. The third major wave was precisely the risen of BPM as a process-centered management discipline, where business processes are perceived worldwide as the backbone of business organizations and their inter-organizational supply chain [12, 54]. BPM has been widely applied across several industries including but not limited to healthcare, manufacturing, and banking [62]. Other significant process-oriented methodologies and studies that have emerged under the BPM umbrella include Statistical Process Control [55], Define, Measure Analyze, Improve, Control (DMAIC) in Lean Six Sigma [69, 70], process-aware organizations [56], and, more recently, process mining, RPA and AI-assisted process automation [17, 71], as well as enterprise process orchestration [72].

2.1.1. Process Modeling and Language Notations

Process modeling in construction is a technique widely used to capture, analyze, control, and improve construction operations. It involves graphically representing construction processes using formal syntax (i.e., modeling notation conventions) and formal semantics (i.e., organized structure with logical branches and meaningful relations). A process model should formally represent the E2E execution of the construction operation, from start to finish, including behavioral state changes, logical sequence, relational dependencies, decision points, involved stakeholders, and constraints [73]. Process models help to unveil, untangle, and study the complexity of construction business operations. They are the foundation of traditional or modern methods such as scheduling, simulation, BPR, BPM, and process mining to study and monitor process performance [74-76].

There exists a broad spectrum of formal modeling language notations to represent business processes. The selection of a given notation depends on the type of process to be represented and on the type of process-oriented analysis to be performed, as each language notation provides specific representation capabilities and comes with some limitations [77]. For instance, Business Process Model and Notation (BPMN) is an international process modeling standard widely used in several industries since it provides formal semantics on process control-flow promoting process transparency and communication among process participants [78]. However, it has the limitation of not being able to represent the spatial locations of process activities [79, 80]. Regardless of the selected modeling language, the structural soundness and the quality of any discovered process model should be evaluated in terms of how well it represents the actual behavior of the system [55]. Table 2-1 provides a detailed comparison of the most commonly used process modeling notations highlighting their main characteristics, capabilities, and limitations.

Table 2-1. Process Modeling Language Notations

Method & Reference	Description	Capabilities	Limitations
Integration Definition (IDEF0) (IDEF3) [81, 82]	Hierarchical functional method containing units of information. IDEF 3 includes object state transition networks.	It describes each activity as a combination of subprocesses. It captures process functions, data, and objects.	It lacks the ability to model control flow of discrete event systems. Non-graphical representation of resources (queues). Complex representation.
Role Activity Diagram	It captures actor roles within the process.	It captures roles, activities, and their interactions.	It lacks the ability to capture the change of states.

[77]			
Unified Modeling Language (UML) [83]	The process diagram represents the execution of a process in machine-readable form.	It captures process structure, behavior, and functionality. It is broadly used to model software systems.	It fails on capturing involved resources.
Discrete event system specification [84]	Mathematical object/system representation.	It represents complex discrete-event systems.	It is perceived as a complex process visualization (i.e., it is not suitable for proper communication).
State Charts [77]	These diagrams represent sequences of states.	It is broadly used to represent dynamic systems.	It does not capture resource interactions.
Activity Cycle Diagram [77]	This technique represents complex interactions of entities in a system.	It is broadly used to represent Stochastic Discrete Event Systems.	It is prone to overfitting and presents difficulties for capturing complex networks logics and activity's locations.
Event-driven Process Chains (e-EPC) [85]	Graphical Business Process Reengineering language.	It represents the flow or sequences of functions and events.	It does not properly capture state changes.
BPMN [78]	International standard graphical language. It is an official standard for the ISO.	It represents the sequence of work and events including major decision points. Simple and functional notation. It captures process concurrency.	Unable to capture activity's spatial locations.
Cell-based modelling [84]	It represents spatial resources in construction	Captures spatial locations. Space is divided into cells as discrete events models.	Difficulty to represent complex site layouts.
Directly-follows graphs (DFG) [11]	It is a graph with nodes representing activities and with directed edges depicting directly-follows relationships.	Useful to represent in a simple manner a sequence of activities happening in the process.	This representation does not usually capture concurrency. Not recommended for advanced process analytics with complex behavior.
Process Trees [11]	Graph notation in the form of a block-structured hierarchical workflow net with top operator nodes and activity leaves.	It helps to ensure the soundness of model representations. It can be transformed into textural form and other process notations.	It focuses on control-flow making it less suited for unstructured processes. Limited basic process modeling constructs.

Petri Nets [86]	This is a mathematical representation of the flow of work as a series of places, arcs, and transitions.	It captures process concurrency. It can make use of colors for distinguishing different process states.	Unable to capture activity's locations. It is prone to overfitting & experience deadlocks.

2.2. Process Mining Foundations

2.2.1. Process Mining

Process mining, a fast-growing field of study and modern digital innovation, has become a crucial asset for business organizations seeking to advance digital transformation across multiple industries [29, 30]. This industry-proven technology enables organizations to leverage real-world project operational data collected from heterogeneous digital sources to automatically discover actual E2E process model behavior and to dynamically assess, monitor, and improve operational performance [31]. In this vein, augmented process transparency, strategic decision-making support, improved process efficiency, and increased productivity are among the most frequently reported benefits of process mining in the literature [87, 88].

Wil van der Aalst (2016) describes process mining as "the missing link between data science and process science" [11]. From one side, this research field falls under the broader umbrella of Business Process Management (BPM), leveraging the latest technological developments to manage and improve business processes. On the other side, process mining algorithms are grounded in data science, as they exploit large volumes of digital records from daily operations to dynamically E2E processes, monitor their actual performance, and enable AI-driven automation capabilities [17]. This novel technology focuses on the exploration of digital traces (e.g., sequences of executed events) through advanced sequence mining algorithms to discover and analyze business process performance. An E2E process model representation includes a clearly defined start, the complete set of sequentially ordered and interrelated activities executed across one or more process instances, and a clearly specified end.

2.2.2. Process Mining Tasks, Tools, and Applications

Knowledge extraction through advanced process mining algorithms provides objective insights into process execution patterns and performance behaviors, which can be leveraged to perform the following process mining tasks: (i) deriving and documenting E2E business process models; (ii) assessing process performance through frequent audits and automated conformance checking to detect bottlenecks and deviations [89, 90] ; and (iii) re-engineering existing or newly derived processes based on action-oriented improvements identified during performance assessment and diagnosis [91, 92], as outlined in Figure 2-1.

Four major types of construction process models were identified in the existing PMM literature [26]: (i) The “as-planned” process models describe the expected or ideal process behavior and serve as the baseline for re-engineering efforts and Business Process Management (BPM) initiatives; (ii) The “as-is” processes refer to the reference models already documented by organizations, which may or may not reflect actual execution. These models are commonly used for internal control, compliance checks, and process audits; (iii) The “as-happened” (or “as-executed”) end-to-end (E2E) processes represent the actual sequence of state changes as they occur in practice. These can be captured in near real time or reconstructed from historical event data stored in databases or information systems; and (iv) The “to-be” process models represent re-engineered or optimized versions of the process, designed based on insights gained from process analysis and exploration of alternative behavioral ‘what-if’ scenarios [93, 94].

The global adoption of process mining has significantly increased in recent years. A variety of software tools, libraries, and platforms exist to support its application, including Disco [95], ProM [96], Celonis [97], PM4Py [98], bupaR [99], PMTk [100], Microsoft Power Automate [101, 102], IBM Process Mining [103], Palantir [104] and Apromore [105], among others. In this regard, Gartner (2023) recently published a global market assessment that evaluated several process-mining platforms in terms of their adoption and functional capabilities for data-driven, process-oriented analytics, with Celonis currently positioned as one of the leading vendors [106].

Process mining has proven to be a game-changer for organizations across various industries by enabling more effective and automated management of business operations. Industries such as retail, manufacturing, healthcare, finance, and logistics have leveraged data-driven process mining to monitor and improve their business processes, support fact-based decision-making, and drive productivity growth. For instance, the banking sector has applied process mining to detect fraudulent actions, streamline financial processes (i.e., payments), and perform risk and mitigation analyses [29, 107].

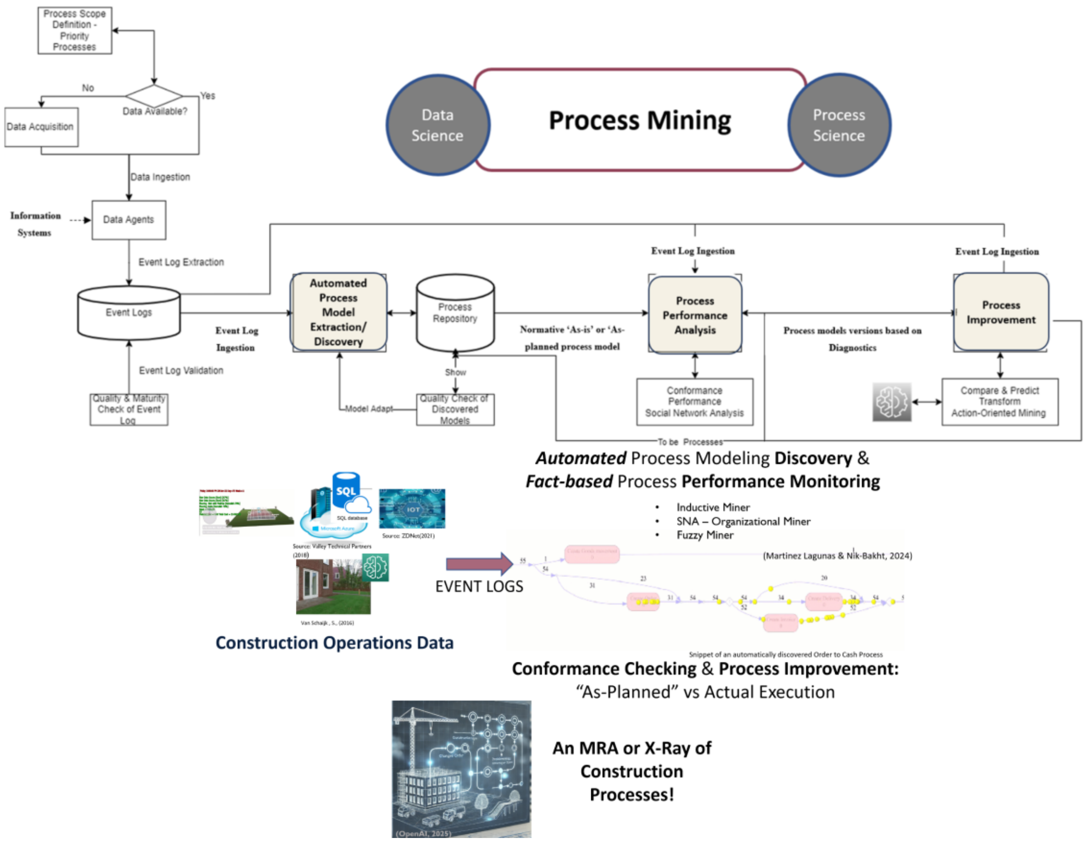


Figure 2-1. Process Mining Overview

Process mining is all about harnessing project operations data by curating this data to extract process-oriented insights that support decision-making in a fact-based and automated manner. In this regard, data curation for the generation of well-structured process event data is an essential and non-trivial task that can take up to 80% of the process mining implementation effort [31, 76, 103].

2.2.3. Event Logs: fundamentals, standards and data schemas

The digital transformation of the construction industry is forcing organizations to adopt new digital technologies to remain at the forefront of innovation and stay competitive in a fast-paced, highly demanding, and constantly evolving market [14]. This shift entails ongoing technology scouting, along with the adoption and implementation of various information systems (IS) to improve the management of construction projects and supply chains, including, for instance, CMIS, Workflow

Management Systems (WfMS), and Enterprise Resource Planning (ERP) systems. These systems record valuable transactional project data from companies' day-to-day operations, i.e., daily Human Machine Interactions (HMI) from construction project stakeholders interacting with those IS to perform certain process tasks. However, construction organizations constantly strive to harness this data to extract actionable insights that support data-driven decisions and enhance process performance. In this context, to enable process mining intelligent capabilities such as automated process extraction and dynamic process monitoring, it is necessary to extract this operations data in the form of well-structured event logs from diverse siloed digital data sources commonly used in construction projects. These data sources can include Structured Query Language (SQL) databases, Workflow Management Systems (WfMS), IoT sensors, BIM, three-dimensional (3D) point clouds, Process-Aware Information Systems (PAIS), CDEs, and Project Management Information Systems (PMIS), ERPs, among many others, most of which are enabled by the so-called 'Internet of Events' [11, 35, 89]

The quality of event logs is of paramount importance to ensure meaningful and successful process mining implementations. Event logs must be machine-readable with well-defined data structures and semantic relationships (i.e., with formal syntax and semantics). To meet these requirements, they must follow existing international event log standards such as the eXtensible Event Stream (XES) standard written in an Extensible data schema definition (XSD), and the most recent Object Centric Event Logs (OCEL) standard, allowing for several flexible data schemas in Extensible Markup Language (XML), JavaScript Object Notation (JSON), SQL relational database structure, or as a graph database using Labeled Property Graphs (LPG) embedded within Event Knowledge Graphs (EKG) [76, 108-110].

Process event logs contain a collection of events with valuable digital footprints or traces of the actual construction process execution histories. An event is composed of at least three base attributes: (i) cases, referring to the process instance containing the collection, group or class of events; (ii) event activities, which are transactions or occurrences happening at a given point in time; and (iii) timestamps of the executed event activities that often follow the ISO 8601 standard [11, 91]. These event logs can be enriched with additional dimensions, such as objects, actors, required resources, and cost, to enable process analysis from multiple perspectives [108, 109].

2.2.3.1. XES Standard

The XES standard was created in 2009 by the Architecture of Information Systems (AIS) research group from the Eindhoven University of Technology (TU/E) [111] and in 2010 was adopted by the Institute of Electrical and Electronics Engineers (IEEE) [112]. Nowadays, it is a well-recognized international standard for structuring, storing, and interchange event logs in a machine-readable representation primarily based on the Extensible Markup Language (XML) suitable for process mining implementation [113]. The XES standard is founded on the concept of a single case notion, meaning that events and their related attributes belong to one process instance. The XES standard is founded on the concept of a single case notion, meaning that events and their related attributes belong to one process instance. CSV tabular event logs can be transformed into XES using ProM [96], XESame [114],OpenXES [115] or PM4Py [98].

In traditional process mining, i.e., with event logs constructed according to XES standard), events are assumed to be partially ordered, meaning that they follow a strict partial order that is transitive (If a≺b and b≺c, then a≺c), irreflexive (a ⊀a), and asymmetric (If a≺b, then b⊀a). An event log can be formally defined by (Eq. 1) [91].

$$EL\ (XES) = (E,\ F_{(attr-val\ map)}, \prec \qquad \text{(Eq. 1)}$$

Where E is a collection or set of events; F (attr-val map) is a mapping function of attribute-value pairs to the corresponding events; and ≺ represents the partial ordering of events often based on timestamps. For control-flow process analysis, this event log notion can be further simplified as per (Eq. 2) [91] given that only the activities' key-value pairs need to be sequentially ordered for a given process workflow instance.

$$EL(XES) = \left(E, \left(\sigma_{(C)}\right), \prec\right) \qquad \text{(Eq. 2)}$$

Where $\sigma_{(C)} = \langle act(e)_{(1)}, act(e)_{(2)}, \dots, act(e)_{(n)} \rangle$ corresponds to a series of traces with the process event activities sequentially ordered for a given case.

Raw process-related data can be transformed into standard process event logs that comply with XES standard. Figure 2-2 shows the UML class diagram for the event log data schema according to XES standard, including its main components such as logs, extensions, global attributes, classifiers, traces, events, attributes, and data types [2, 108, 116]

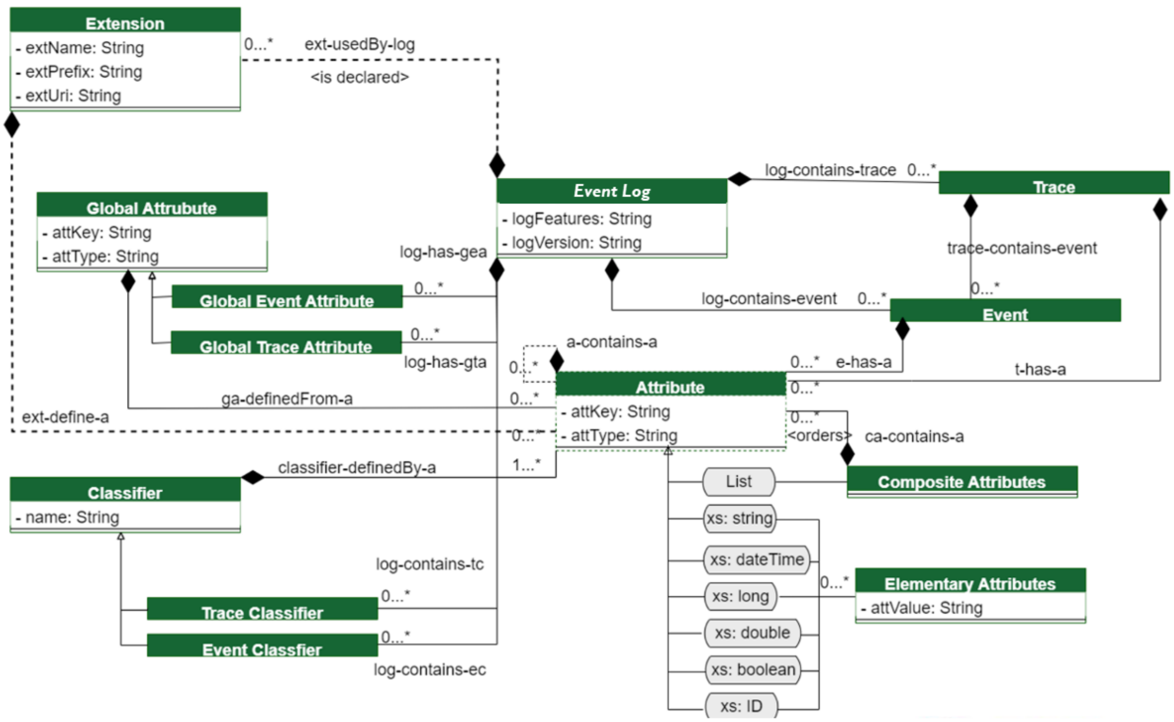


Figure 2-2. XES UML class diagram, modified from [108]

Event Log – An event log refers to a historical record of activities or transactions happening across processes' lifecycle. Events are labeled with execution timestamps and grouped by a unique case identifier (i.e., process instance or trace). An event log can contain one or several event traces, and it should declare/contain any required extensions to semantically describe the process.

*Extensions –*The extensions provide formal semantics and structure to the event log components by considering/assigning predefined attributes at various levels of the event log (i.e., log, trace, event). There are seven standard extensions in the 2016 version of the XES standard (IEEE, 2016). New custom extensions can be defined for domain-specific event log developments.

Global Attributes – This class refers to the declaration of global attributes used when certain predefined process-related information needs to be contained in the log. These global attributes are assigned to every trace and event within the log.

Classifiers – An identity can be assigned to each trace and event in the whole log using classifiers. These classifiers act as labels for traces and events that allow grouping them to compare against one another. An example of an event classifier shall contain two main attributes the event activity/instance name and the lifecycle transition of that activity (i.e., "Create Event – Complete").

Traces – They store the event activities related to a process instance or case. Each trace can contain several event objects. A trace represents a series of activities during the execution of a given process instance; these activities are partially ordered based on timestamps.

2.2.3.2. OCEL Standard

XES standard is reliant and limited by the single case notion assumption, which states that a process event belongs to a single case (i.e., process instance), activity and timestamp. XES is useful to analyze a particular process from a single process perspective, but for real-world complex processes such as those in construction projects, process events can be interrelated to multiple intertwined processes, and certain process event activities may refer to different process entities. The XES schema, thus, does not to capture this multidimensionality and intertwinement of real-world processes. Considering this aspect, the Object-Centric Event Logs (OCEL) standard aims to overcome this limitation [117].

The OCEL 1.0 standard, developed by the Architecture of Information Systems research group in 2020, is an object-centric even log structure that empowers business analysts with the capability of structuring and storing multidimensional business processes. OCEL provides a more realistic view of business process behavior and normally stands between the source data coming from IS and the XES event log extraction. Multidimensional OCEL logs can be flattened into XES logs as needed for further analysis on specific process perspectives/views, yet one should be aware that the flatting approach can result into discovering false process behavior due to convergence (i.e., duplicated events) and divergence (i.e., considering or omitting events that are not part of the selected perspective) problems [118].

Aiming to improve and simplify OCEL 1.0, OCEL 2.0 standard has been recently released to facilitate the schema definition of multidimensional event logs in the form of a federated Object Centric Event Data Models (OCDM). Figure 2-3 provides a visual representation of OCEL schema definition in a UML class diagram. The main OCEL 2.0 components are the log, events, objects, event types, object types, event-to-object relationships, object-to-object relationships, and their related attribute-value pairs as per (Eq. 3) [117]. OCEL 2.0 data structure also sets aside the global classes previously considered in OCEL 1.0.

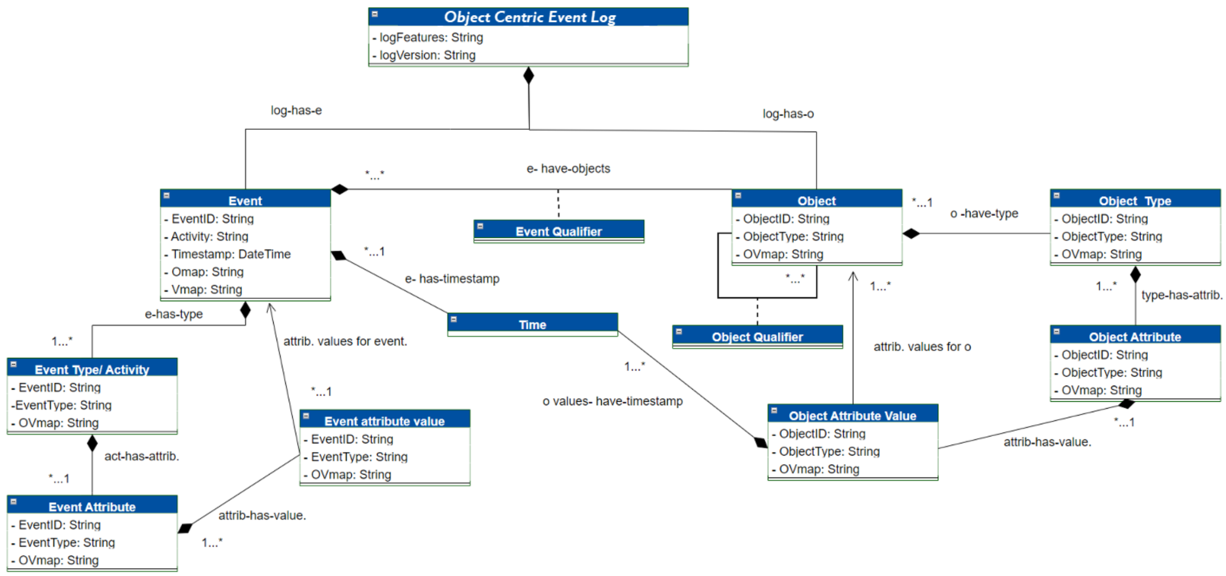


Figure 2-3. OCEL 2.0 UML class diagram, modified from [117]

<u>*OCEL Log Class–*</u> Similar to XES, but without the trace and extensions class definitions.

<u>*Objects–*</u> This is the most important component of Object Centric Event Logs. Instead of traces for single case notions. An object is any business operational entity that participates in or is affected by event activities during process execution.

<u>*Object Type–*</u> The category or class of object entity. OCEL is composed of a list of object types that can be seen/defined as multi-case event logs to discover a multidimensional process model. Each object type can have multiple object instances. OCEL event logs can be flattened to XES logs via the objects.

<u>*OCEL Events–*</u> Similar to XES, but they are related to objects rather than to traces.

<u>*Event-to-object relationships–*</u> Each event can be related to one or more objects.

<u>*Object-to-object relationships–*</u> Objects can be related to other objects.

<u>*OCEL Attributes–*</u> Similar to XES, but related to objects and events (i.e., not to traces).

OCDM can be generated based on OCEL standard to allow for multi-perspective process analysis where a process event can be related to any number of process entities known as 'objects. For OCEL, it is assumed that objects and events should be pairwise disjoint, meaning that objects can not be represented as events, and that a total ordering based on timestamps is considered. OCEL can be formally defined as a tuple according to [117]. Multidimensional event logs constructed in CSV tabular can be converted into JSON OCEL using the Object Centric Process Analysis (OCPA) python library [119]. OCDM can also be stored in graph databases as EKG to provide enhanced semantics and filtering capabilities across E2E process components [120].

$$OCEL = (E, O, EA, OA, evtype, time, objtype, eatype, oatype, eaval, oaval, E2O, O2O) \quad \text{(Eq. 3)}$$

Where E, O, EA, OA correspond to set of events, objects, event attributes, and object attributes respectively; time, refers to the event timestamps; objtype, indicates the type of object; eatype and oatype, assign types to event and object attributes respectively; eaval and oaval, refer to the assigned values for events and object attributes; lastly, E20 and O2O, consider the relationships between events to objects and objects to objects respectively [91]. An aggregated comparison of both XES and OCEL standards highlighting their identified limitations and main capabilities is presented in Table 2-2.

Table 2-2. Event Log Data Standards Comparison Overview

Criteria	XES [108]	OCEL [117]
Perspective	Perspective of interest	Multi-perspective
Approach	Traditional Process Mining	Multidimensional Process Mining
Quality Aspects-Transparency	Less transparency into events and precedence relationships	Preserves entities/objects, events & relationships – Higher transparency
Suitable for	Single workflow executions	Interrelated E2E processes
Case concept	Single-case per trace	Multi-object, object-centric
Components	Log, traces, events, relationships	Logs, objects, events, relationships
Ordering Type	Partial order (i.e., subsets-based)	Total order (fully timestamp-based w/o subsets)
Conversion Aspects	XES to OCEL (Requires domain knowledge)	OCEL to XES (check for convergence and divergence problems)
Process Representation	E2E 2D process model	EKG/ Advanced Process Model
Interoperability	Enables Conformance Checking	Enables Conformance Checking
Conformance	Less focused on interoperability	Increased interoperability

2.2.4. Process Mining Perspectives

Under the XES standard, event logs can be constructed and structured according to the perspective of interest to enable the discovery of E2E process behavior and the analysis of process performance. In practice, five main perspectives are commonly examined by organizations: the E2E control-flow perspective (how activities are sequenced and how the process flows), the organizational perspective (resource involvement and capacity), the case perspective (behavior at the individual process-instance level), the cost perspective, and the time perspective [11]. Other custom perspectives can be added to XES event logs for further enriched process-oriented analysis [2].

2.2.5. Process Discovery Algorithms

The α-algorithm was one of the first formal process mining techniques developed to discover real-world process model behavior. It works by scanning the event log to derive a footprint matrix that captures the ordering relations or sequence patterns observed between each pair of process activities within the event log, which are also known as transitions. A process model is then extracted or discovered based on the footprint matrix [11]. Although this technique supports concurrent process behavior, it infers this behavior rather than explicitly detecting it. Besides, it lacks the capability to accurately handle process loops, specifically when the same activity sequences within a trace are executed multiple times in succession. Consequently, the process models derived through this technique tend to be overly complex and unrealistic as they present dense, spaghetti-like structures that lack interpretability [11]. Although the α-algorithm often proves ineffective due to its inability to filter out non-significant behavior, it laid the foundation for more advanced process discovery techniques, such as the Fuzzy Miner and the Inductive Miner (IM) algorithms [121-123]. These advanced algorithms are widely used in real-world industrial applications, as they support the extraction of more realistic and interpretable process models by filtering out non-significant execution paths (i.e., Fuzzy Miner) and ensuring soundness or the absence of deadlocks (i.e., IM), which allows the E2E execution of each process instance.

The Fuzzy Miner algorithm provides a simplified aggregated view of a process with high interpretability by constructing hierarchical Directly-Follows Graphs (DGF) from analyzing each pair of subsequent process activities, but unlike the α-algorithm, the Fuzzy Miner reduces process complexity by hierarchically ordering the process behavior based on two metrics: the frequency-based activity significance, and the closeness-based correlation that help determine the connection strength between a pair of subsequent activities. In other words, the proximity correlation determines the connector thickness between two process steps on the Directly-Follows Graph (DFG) based on how often the two events occur in temporal succession within the event log [121, 124]. In a nutshell, this algorithm helps to (i) preserve frequent process behavior, (ii) cluster and aggregate less frequent but highly correlated process activities, and (iii) filter out infrequent activities and weakly correlated relations.

Fuzzy Miner DFG can be constructed using (Eq. 4) [121], where EA' and ER' are the subsets of event activities and relations in the event log (L) whose relative significance (σ) and proximity correlations [$corr(ea_1,ea_2)$] are greater than the defined significance (θσ) and correlation (θ_{corr}) thresholds. A common approach to define these abstraction thresholds is based on the Pareto principle, which establishes that keeping 20% of the most frequent activities and their relations often cover 80% of the observed process behavior [125]. However, despite of its simplicity and high interpretability to discover unstructured processes, the Fuzzy Miner algorithm fail to detect concurrent behavior and it is prone to producing unsound imprecise process models that does not accurately represent the observed behavior a result of its filtering mechanism [124].

$$DFG(L) = (EA', ER'); \quad \text{(Eq. 4)}$$
$$\forall ea_i \in EA \wedge \forall\ (ea_1, ea_2) \in ER;$$
$$\sigma(ea_i) \geq \theta\sigma \rightarrow ea_i \in EA' \wedge corr(ea_1, ea_2) \geq \theta_{corr} \rightarrow (ea_1, ea_2) \in ER'$$

To overcome these limitations, the IM algorithm adopts a hierarchical process modeling notation whose inner nodes are block-structured execution operators, including sequence (→), parallel (∧), exclusive choice (×) and redo loops (⟳) relational operators; and its tree leaves are process activities executed according to the type of operator. Unlike other process discovery algorithms, the IM and its tree-like notation guarantee process modeling soundness while handling infrequent behavior through well-structured and simple semantics. As described by (Eq. 5) [126], the IM algorithm, rooted in DFG, aims at discovering process trees by recursively splitting the main event log into sublogs, where each sublog can be represented into a subgraph or subtree based on the cut type [123]. The recursive cuts depend on the type of relational execution operators observed in the log or sublog. The main event log is repeatedly partitioned until single activity sublogs are derived, meaning that no more cuts are possible, and leaf nodes are reached [122].

$$Q(L) = (\oplus, \mathrm{QL}' \mid_{A\,start}, \mathrm{QL}' \mid_{Ai \ldots An}, \mathrm{QL}' \mid_{A\,end}); \quad \text{(Eq. 5)}$$
$$\oplus \in [\rightarrow, \times, \wedge, \circlearrowright]$$

Where the process tree (Q) of the main event log (L) is constructed by partitioning ($\oplus$) the multiset of traces contained in the log into subtrees (QL') based on the observed relational execution operators. These subtrees are normally constructed from the start sublog ($\mathrm{QL}' \mid_{A\,start}$), the end sublog ($\mathrm{QL}' \mid_{A\,end}$); and the intermediate sublogs decomposed based on the directly-follows relationship of event activities observed in its traces. Process trees derived through the IM algorithm can be converted to other process modeling language notations such as Petri Nets and Business Process Modeling Notation (BPMN) as needed for the documentation and analysis of the process [11, 127]

CHAPTER 3. LITERATURE REVIEW[2]

Unveiling and analyzing E2E process behavior is essential for diagnosing operational inefficiencies and identifying opportunities for targeted process improvement and automation. [93, 128]. Thus, this chapter dives deeper into the topic by providing an exhaustive review of the current state of the art on process mining, modeling, and management in the AEC/FM industry, analyzing these areas both individually and holistically as interrelated fields of study. The analysis covers process-oriented studies in the construction domain published in the last three decades. The ultimate goal of this in-depth literature analysis is to lay the foundation for the design and development of a process-oriented framework tailored to the construction domain focused on automated modeling and dynamic performance monitoring of construction processes. To this end, the specific objectives of this analysis include: (1) identifying existing PMM methods, current practices, and main understudied research areas; (2) identifying critical construction processes across the entire lifecycle of construction projects; (3) spotting research gaps and future directions; and (4) designing a process-oriented tool to support construction practitioners in managing, improving, and automating construction operations.

3.1. Literature Research Method

Construction projects are process-heavy. Unlike processes in other industries, construction processes are frequently characterized by heterogeneous and non-standard practices, scarce process control, and low transparency and visibility. Consequently, critical processes can significantly endanger the overall performance of construction projects if they are ineffectively managed or inefficiently executed [27]. Hence, analyzing the existing literature is essential to investigate current PMM practices. The systematic method used to survey the existing PMM literature, shown in Figure 3-1, followed two main phases: (i) definition of scope, comprising data collection, screening, and classification; and (ii) in-depth literature analysis, including scientometric and systematic reviews. As a result, three major outcomes were produced: (1) a

[2] This chapter is based upon:
Martinez Lagunas, A. J., & Nik-Bakht, M. (2024c). Process Mining, Modeling, and Management in Construction: A Critical Review of Three Decades of Research Coupled with a Current Industry Perspective. Journal of Construction Engineering and Management, 150(11). https://doi.org/10.1061/jcemd4.coeng-1472

conceptual PMM classification framework of analyzed process-oriented studies, along with the identified major research gaps; (2) a portfolio model of the core construction processes reported in the literature across the entire construction project lifecycle; and (3) a process management canvas tool designed for construction practice.

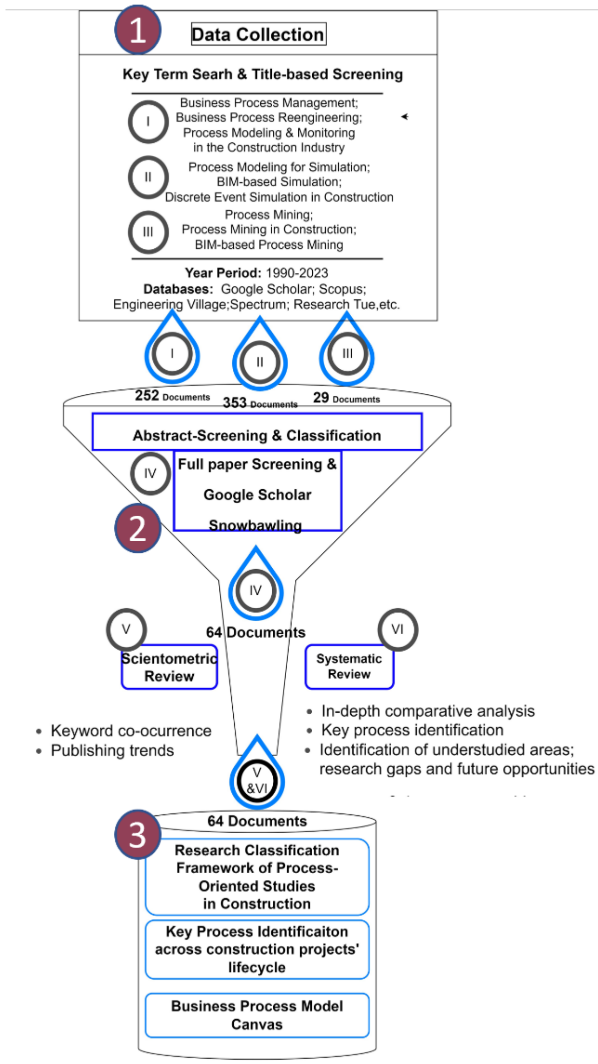


Figure 3-1. High-level literature surveying method: data collection, literature review and outcomes.

Process-oriented studies were thoroughly analyzed as reported by the most current state of the literature to identify and highlight understudied research areas, publishing trends, core processes in construction projects, means and methods driven by ICT, and major gaps.

3.1.1. Literature Data Collection

The data collection phase consisted of searching for and gathering process-oriented studies from published academic theses and topnotch journals such as the Journal of Construction Engineering and Management, Automation in Construction, Advanced Engineering Informatics, among others. The data sources for this collection included Google Scholar; university research repositories such as Spectrum, TU/e, and TSpace; and technical databases such as Engineering Village, Scopus, and Web of Science (among others).

Since the study of business processes is not exclusive to the construction domain and inherits several aspects from manufacturing and other industries, the authors performed search queries on three major levels from general to specific. The first level included search queries such as: {"business process management" OR "BPM" OR "BPM in the construction industry"}; {"business process reengineering" OR "BPR" OR "BPR in the construction industry" OR "business process redesign" AND "process improvement" OR "process optimization" OR "process automation"}; and {"process modeling" OR "process modeling in the construction industry" OR "process mapping" OR "process monitoring"}. The second level included the following search query: {"process modeling for simulation" OR "simulation in construction" AND "construction operations" OR "BIM-based simulation"}. Finally, the third-level search query was as follows: {"Process mining" OR "Process mining in the construction industry" OR "BIM-based process mining"}. These search queries resulted in a comprehensive set of 634 documents comprising technical journal articles, books, and theses. Given the large number of documents, two main exclusion criteria were defined to set aside two groups of publications that require a separate study: (1) the body of work on activity recognition; and (2) the rich body of knowledge on construction simulation. The former, while an essential component of automated process modeling and management (PMM), was set aside since this study scopes "E2E processes" (rather than individual activities). From the latter category, while this research took advantage of studies that simulate an E2E construction process (i.e., PMM for simulation), a considerable number of publications were not used in this research, as the analysis of construction simulation literature calls for a separate study.

3.1.2. Literature Screening and Classification

After defining the search criteria as shown in Fig. 1, the second and third phases included title and abstract screenings that helped filter and classify the literature collection into 98, 64, and 18 studies corresponding to each major category: (1) BPM and process modeling; (2) process modeling for simulation; and (3) process mining, respectively. Then a full article screening was performed resulting in 64 documents for the three categories. Ultimately, a "snowballing" technique was implemented consisting of reviewing reference sections from each document to find any additional relevant studies on Google Scholar through their corresponding citations. The final stage consisted of an in-depth comparative analysis that scrutinized more than 96 documents to extract detailed and aggregated insights from them. Given that most process-oriented studies are indirectly or directly related to BPM, the main selection criteria centered on the BPM life-cycle phases: process identification, discovery, modeling, analysis, redesign or re-engineering, implementation, monitoring/controlling, and improvement. Thus, if the studies covered one or more of these phases with a clear distinction of the studied process, methods, analyzed components and characteristics across any of the major project life-cycle stages, they were selected, fully scrutinized, and aggregated as described in the following section.

3.1.3. Keyword Co-Occurrence

Keyword co-occurrence is a quantitative technique that falls under the umbrella of scientometric research, which aims to identify main scientific themes, network analysis, trends, and other research patterns from the entire collected body of knowledge. In this research, keyword cooccurrence served to perform a bibliometric network analysis of the collected process-oriented literature by quantitatively identifying pairs of keywords that frequently occur in the literature.

This frequency-based analysis helped to identify major scientific research knowledge areas and their most frequent interrelations [129]. As a result, a network composed of four major interrelated clusters was found by analyzing the co-occurrence of key terms frequently contained in the collected literature. Figure 3-2 shows the identified major clusters: process mining; process modeling; process management interlinked by a fourth central cluster, construction management. The sphere node sizes show the overall importance degree centrality of each term, meaning that nodes holding more edges are of higher relevance, and the thickness of arcs/edges indicates the strength of interactions between connected terms based on their frequency of occurrence in the literature.

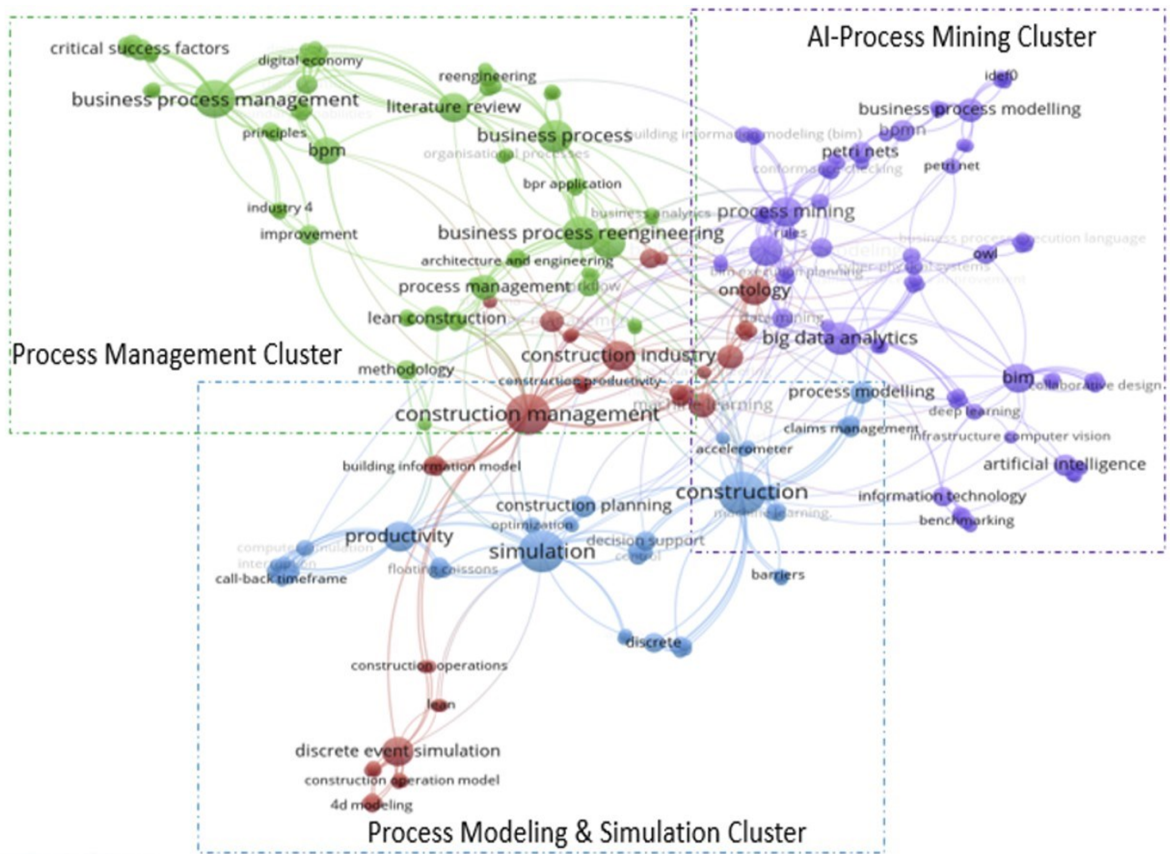


Figure 3-2. Process-oriented clusters in the PMM literature

3.1.4. Publishing Trends

As part of the scientometric analysis, the number of published process-oriented studies per year was investigated, as aggregated in Figure 3-3**.** It is worth noting that the bulk of the reviewed studies (73 out of 96) were published in the last three decades. Also, Figure 3-3 provides a drill-down view of these publications per year per cluster. With the advent of the internet and cutting-edge ICT developments, the number of publications significantly jumped in 2010, keeping an exponential growth to date, including studies on BPM, formal knowledge process modeling, process modeling for simulation, and process mining. Another relevant aspect to highlight from the Figure 3-3 is that studies published before 2008 mainly belong to the process re-engineering, management, and modeling clusters setting the knowledge foundation for more advanced studies [68, 130].

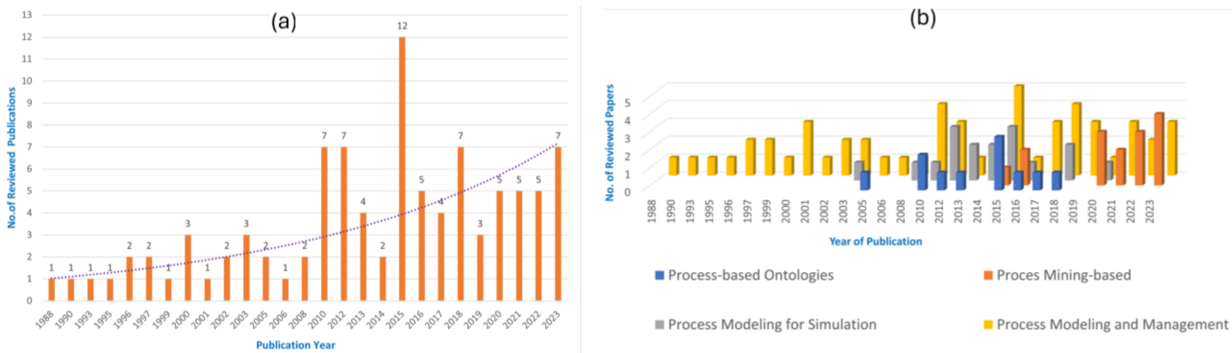


Figure 3-3. Publishing trends on process modeling and management (PMM): (a) publications over the years; and (b) publications

3.2. PMM in Construction

Construction projects comprise a broader scope than construction processes and must be completed within agreed deadlines. Yet, projects can be viewed as a compendium of major processes (i.e., work packages) required across the entire project lifecycle, from conception to realization, and involving the execution of several interrelated subprocesses [131]. However, unlike construction projects' unique, dynamic, and temporary nature, construction business processes are expected to be reusable, customizable, and interoperable across similar projects to ensure their technical compliance, meet quality requirements, and streamline construction operations. Therefore, process complexity levels vary according to projects' specificities and stakeholders' roles and responsibilities in those processes [83].

Effective and efficient process management across a construction project's life cycle is crucial not only for contributing to the project's success but also for overcoming the productivity challenge faced by the construction industry. In this regard, Dixit, et al., (2019) reported several factors that impact construction productivity involving drawing errors; labor, material, and equipment unavailability; untimely communication of information requirements; and poor coordination and management between parties. Along the same line, from a meta-analysis of 26 studies, Horman and Kenley (2005) concluded that 49.6% of construction operative time is wasted in non-value-adding activities [132].

The construction sector recognizes the need for a structured process-oriented knowledge management approach to optimize operations and mitigate the negative impacts on construction productivity [128, 133-135]. Addressing this need can result in significant quality improvements, cost savings, and on-time project delivery. Sanvido (1988) introduced a conceptual construction process model (CCPM) that focused on control roles at various organizational levels to assist construction forepersons. Implementing CCPM led to enhancements in productivity and budget savings of up to 5%. CCPM aimed at streamlining construction process management [81].

To further analyze construction processes and their effect on organizational performance, it is crucial to identify major processes' components, their relations, and characteristics. In this regard, El-Gohary and El-Diraby (2010) provided a comprehensive process knowledge representation for the construction domain including process activities, process perspectives, task dependencies, and rule constraints [133]. Moreover, Nik-Bakht and El-Diraby (2015) contributed to the body of knowledge by conducting an extensive review on decision-making in the construction domain [136], highlighting the importance of Social Network Analysis (SNA) in uncovering hidden organizational performance patterns in construction projects and how they influence decision-making [137].

From the reviewed literature, and to obtain a high-level view of the significance and complexity of construction process management, I developed the high-level construction process taxonomy shown in Figure 3-4, highlighting major components organized in classes and subclasses that in turn were derived from commonly used industry terminology of the main intrinsic and extrinsic process characteristics. This figure was derived from the systematic method depicted in Figure 3-1 in a way that selected studies were categorized in four major clusters during the scientometric review as shown in Figure 3-3 (i.e., ontology, process-mining, simulation, and general PMM). Then, based on this categorization, the "As-planned" element under the "process model types/states" component depicted in the figure was derived from reviewing ontology-based studies [24, 133]; the "As-is and As-happened" element, from process-mining and PMM studies [93, 138-140]; and the "To-be" element from simulation-based studies [141, 142], to mention just a few. Subsequently, the "Knowledge Assets" and "Constraints" components [53, 141, 143] were inspired by Golzarpoor (2017), Hosny et al. (2020), and Liu et al. (2015); the "Nature" component [94] was derived from was derived from Abourizk et al. (2016), and all other components and elements were derived from the aggregated review analysis summarized in Table 3-1 and Table 3-2. As depicted in Figure 3-4, the main challenge lies in managing all of these components, their complex network-like interactions, and their influence on process performance. The complexity of construction processes varies according to the type and size of projects. These components can be part of a local or global supply chain comprising several project stakeholders interacting to deliver a valuable product or to provide a high-quality service to clients and end-users. For this to happen effectively, it is essential to consider the dynamic nature of construction processes and the required efforts to integrally manage a project [144], including among others (1) effective allocation of resources and organization of work for achieving on-time and on-budget project deliveries; (2) identification of stakeholders and interface coordination; (3) proper communication, information exchange, knowledge management, and collaboration among teams; (4) consideration of process constraints, states, and performance; (5) identification of any potential risks along with their mitigation; and (6) waste identification and management across the entire project life cycle [58, 145].

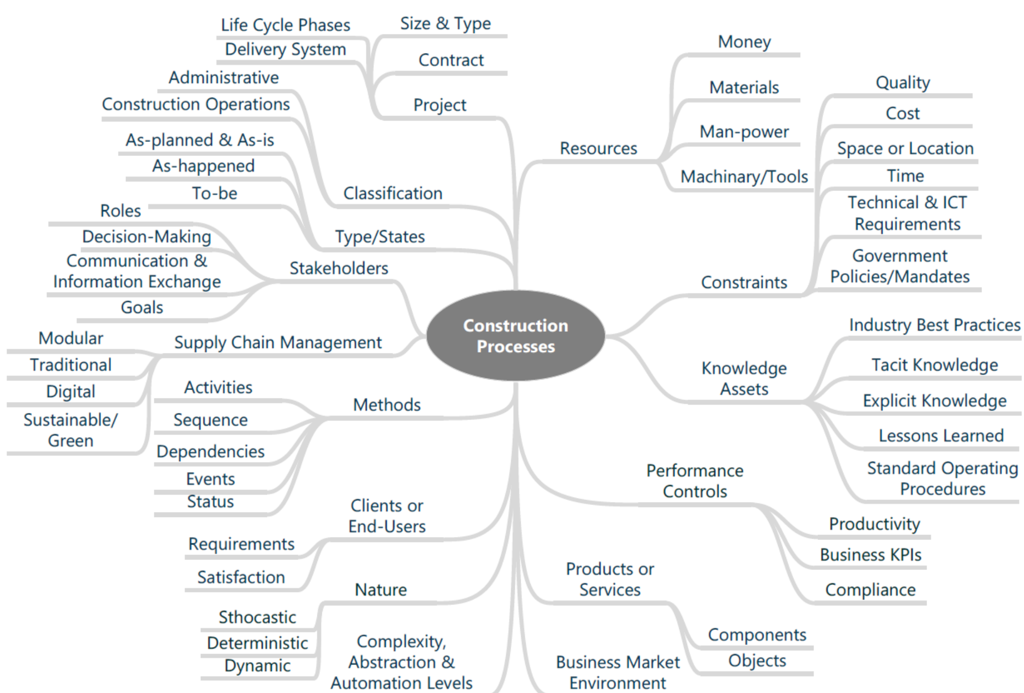


Figure 3-4. Construction process taxonomy - components and characteristics

3.2.1. Process Modeling in Construction Simulation

Conceptual process modeling has been widely used in the construction industry for scheduling and construction simulation since 1957 and 1970, respectively [74, 146-148]. Construction simulation is described as a mathematical representation of building systems by computer software, where processes are modeled using graphical interrelated entities and coded algorithms to mimic the dynamic real system behavior so that it can be improved based on what-if analyses of multiple execution scenarios [94]. In this study, although the scope does not include reviewing the rich domain of simulation in construction, it is essential to look at this domain from a PMM perspective as one of the main areas where contributions to identification and modeling of construction processes are reported. The focus, however, is on how the construction operations models are derived and represented, their main capabilities, and limitations against other PMM studies, and the way their model performance is measured and monitored (rather than the simulation technique used).

Although discrete event simulation (DES) considers the stochastic nature of construction processes, it turns out that most existing DES-oriented studies still heavily rely on manual and subjective methods for collecting data and modeling construction operations [142, 147, 149]. Besides, it is worth noting that DES alone does not suffice to accurately capture and study the complex and dynamic system behavior of construction processes. Thus, recent studies have shifted toward more hybrid methodological approaches that integrate DES with system dynamics (SD) or with agent-based simulation (ABS) to capture not only stochastic behavior from discrete variables at the operational level, but also dynamic behavior from continuous influencing variables (i.e., context level) that tend to change over time and that should loop back to the operational level to effectively support decision-making [150], as well as collective emergent behavior that results from complex interactions among autonomous agents (i.e., process stakeholders) [151, 152]. However, most of these simulations (inside and outside the construction domain) run on top of "as-planned" processes rather than on top of actual executed "As-happen(ed)" E2E processes which result in a lack of fact-based insights [153].

The integration of hybrid simulation approaches and automated process model discovery through process mining algorithms has tremendous potential for efficiently modeling, managing, and improving construction processes [154]. Process mining provides automated modeling and monitoring capabilities of actual construction operations while hybrid construction simulations provide forward-looking predictions—that is, the "to-be" process analysis based on data-driven knowledge pattern discovery [153, 155-157]. In this respect, various studies have highlighted the importance of process modeling and process mining to develop reliable and robust simulations [84, 158, 159]. Table S1 provides a comprehensive aggregated analysis of the reviewed studies on process modeling for simulation.

3.2.2. Classification Framework for PMM Literature in the AEC/FM Domain

Golzarpoor (2017) reported two principal approaches to derive and study construction processes: bottom-up and top-down. The former represents process models by identifying a core sequential structure from previously implemented construction workflows, the latter derives them from industry best practices generally through the development of domain ontologies, which can be described as formal domain representations that include concepts, actors with clear roles, relations, and meaningful logical axioms or rules [53, 133, 160]. Similarly, besides stressing the importance of managing construction process complexity in achieving project success as depicted in Figure 3-4, the conceptual framework shown in Figure 3-5 is proposed, which results from distilling the reviewed PMM literature and from observing that most process-oriented studies in the AEC/FM domain fall into two major categories according to their process modeling and analysis approaches. On the one hand, the top-down approach is characterized by the generation and analysis of governance process models, best practices, or Standard Operating Procedures (SOPs). On the other hand, the bottom-up approach majorly focuses on deriving and monitoring actual process execution. From the aggregated analysis, a need for a more hybrid approach was identified, which can help to operationalize, monitor, and improve construction processes and their related governance strategies. To this end, a continuous process improvement philosophy was incorporated into the integrated framework [161] as depicted in Figure 3-5. The "Results" section provides further details on the in-depth literature analysis related to the proposed PMM framework.

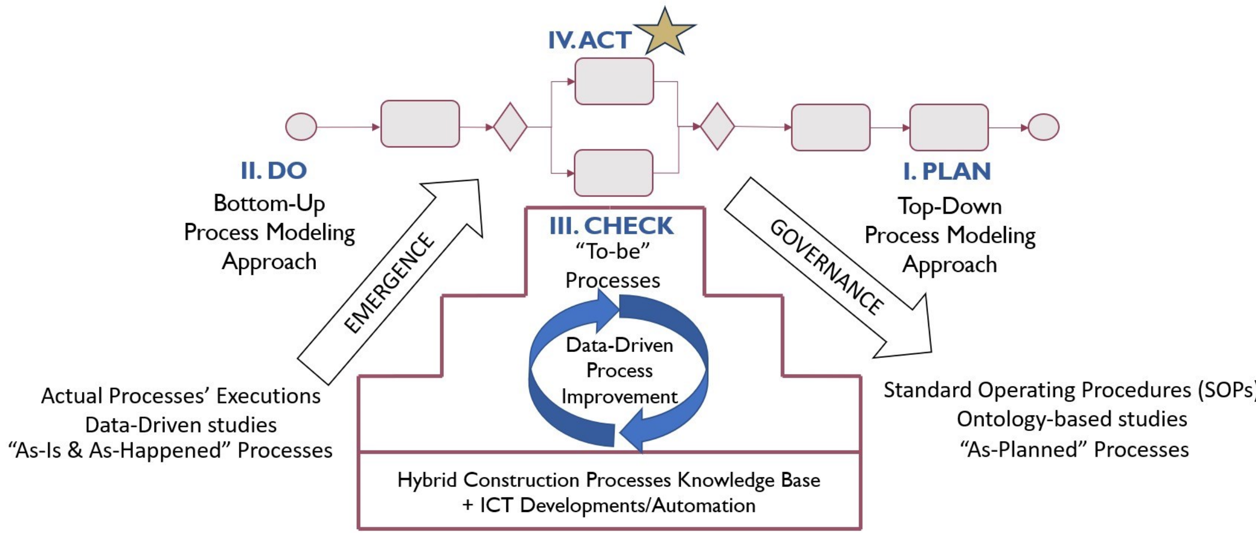


Figure 3-5. PMM conceptual classification framework

3.2.3. Construction Processes in PMM studies

The first step toward construction process improvement and process automation is identifying relevant processes that need to be audited, controlled, or enhanced. Once identified, it is critical to constantly measure and monitor their performance against key performance indicators (KPIs). Thus, business organizations should not only have well-defined KPIs but also have a properly established process classification framework, process portfolio, and SOPs [162]. In this regard, Benevolenskiy (2015) suggested two major construction process categories: (1) administrative, which includes any management and support processes (e.g., progress payment, purchase order approval, request for information); and (2) operational, which involves any core processes related to construction operations (e.g., mobilization, earth moving, steel framing, concrete formwork, finishing) [83]. The former category can be considered service-based processes that support mainstream construction work. They mostly include management and business-centric tasks. The latter category can be considered product-based processes, which comprise mainstream work that takes place in the field and on job sites. An aggregated view of PMM studies within the AEC/FM domain is presented in Table S2 and Table **S3**. This view maps each of the reviewed PMM studies to the construction processes examined in the study and to the associated process model types (i.e., "as-planned," "as-happened," and "to-be") for both service-based administrative and product-based operational processes.

3.2.4. In-Depth Literature Analysis of PMM Studies

The proposed classification framework for PMM studies introduced in Figure 3-5, helps to categorize process-oriented literature from the AEC/FM domain into two major classes according to the process modeling approach adopted in the studies. One is the top-down approach if the process model is considered a priori to the executed actual process—that is, the activities are expected or forced to follow a predefined process sequence (i.e., SOP workflows or governance models). In this case, the processes are usually regulated and imposed by the upper-level management of an organization. Ontologies are frequently developed to build these process models and used as a tool to capture and manage knowledge. Oppositely, if the study aims to “learn” the processes from observing and analyzing data of various activities and events, it follows the bottom-up approach. In this category, predetermined process models are not essential and are expected to “emerge” from the analysis of as-happened operation processes. The latter approach makes use of data-driven techniques, including data mining and more specifically process mining with automated data acquisition technologies to derive these models [163].

3.2.4.1. Top-down PMM (Knowledge-Driven) Studies

Cheng and Tsai (2003) proposed a method for re-engineering construction management processes focusing on improving efficiency by identifying deficiencies before implementing process automation. The proposed method comprised (1) process representation to identify and prioritize core processes; (2) process transformation to model them using integrated definition (IDEF0) notation; (3) process analysis based on a “target achievement matrix” to identify major process defects against target KPIs to evaluate process performance; and (4) process redesign to evaluate process bottlenecks, set redesign rules, and propose process improvements [164]. Even though Cheng and Tsai’s technique to evaluate business processes was useful, it had the following limitations. First, no practical implementation of the re-engineered process was carried out to actively monitor its performance. Second, after identifying major process pain points, no automation strategies were implemented to make the process more efficient (Dave 2017). Third, the study was customer-oriented, but it did not consider industry best practices or lessons learned from previous projects [53].

Expanding on prior work, Cheng et al. (2015) proposed a matrix organization process re-engineering approach that combined business process re-engineering with knowledge management (KM). Core processes were modeled using an event-driven process chain language notation (e-EPC) to evaluate their performance based on KPIs and identification of bottlenecks. Also, performance and service gaps were identified during the process analysis, which set the baseline for process redesign and control. Lastly, reengineered processes were validated against no re-engineered processes (i.e., previous versions) based on a process value gain metric. The authors considered a crucial KM component that allowed organizations not only to analyze and redesign processes but also to retain operational knowledge [165]. However, Cheng et al. did not address automated process modeling techniques, process configuration procedures, or process reusability potentials [83].

Costa et al. (2019) proposed a comprehensive governance framework to manage processes involved in the execution of public construction projects at the organizational level. This framework incorporated a BPM approach comprising four phases: (1) planning and development of organizational business maps; (2) process analysis, consisting of the identification of main process deficiencies and process classification [166]; (3) process modeling of high-priority processes; and (4) processes optimization based on expert knowledge [75]. Despite its broader BPM scope, Costa et al. adopted a manual as-planned modeling approach and subjective validation. Table 3-1 provides an aggregated comparative analysis of major top-down process-oriented studies in the AEC/FM domain, highlighting their main process of interest, the focus of the study, data sources, tools, and methods.

Table 3-1. Main PMM studies – 'top-down' approach

References	Modeled Process	Data Source Details for Modeling Processes	Method or Instrument	Focus of the Study
Costa et al., 2019	Project Delivery Management	Expert's Knowledge	BPM Cycle; Focus Groups	BPM
Golzarpoor et al., 2016	Request For Information (RFI)	Consultation with Experts (Industry Best Practices)	Workflow Inheritance	Conformance & Customization
Golzarpoor et al., 2018	Change Requests and RFI	Reference Models	.NET Windows Communication Foundation. Desktop-based WMS	Interoperability & Customization

Cheng & Tsai, 2003	Procurement/ Purchasing	Expert's interviews & Firm's knowledge	Target Achievement Matrix, Customer-oriented	BPR
Cheng et al., 2015			Target Achievement Matrix + Knowledge Component	
El-Gohary & El-Diraby, 2010b	General	Previous Initiatives & Ontologies	Competency Questions-Protégé-OWL	Transparency, Integration & Reusability
El-Gohary & El-Diraby, 2010a	Urban transportation planning processes	Project Information and previous ontologies	Process Integrator Portal	Interoperability/ Collaboration
El-Diraby et al., 2005	Knowledge Management Taxonomy	Previous Initiatives & Project Documents	Competency Questions & e-CKMI Portal	Knowledge Repository & Transparency
Benevolenskiy, 2015	Produce Precast Columns & Slabs	Previous Ontologies & Existing Standards	Process Configurator-OWL	Dynamic Process Configuration (Rule-based) and reusability

In a high labor turnover industry, preserving construction knowledge is crucial to reduce employee replacement costs and prevent knowledge loss from experienced and skilled personnel [89, 167]. Furthermore, capturing construction industry best practices as a structured process knowledge representation is challenging [53]. In this regard, Golzarpoor et al. (2016) introduced the industry foundation processes (IFPs) ontology inspired by industry foundation classes (IFCs) [168]. The IFP ontology provides a standard structure built on process-aware information systems (ISs) that facilitates process conformance checking to detect deviations between the as-planned governing models (i.e., IFP) and the actual processes. Construction organizations can prevent cost overruns, comply with normative auditory process standards, and increase process productivity and quality by detecting these deviations [24]. The IFP ontology also enables process interoperability across inter-company workflows [169]. However, in comparison with other ontologies in the AEC/FM domain [133], its process classes are far more limited, meaning that additional complex process interactions must be manually configured. Table S4 presents a detailed comparative analysis of reviewed process-oriented ontologies.

El-Gohary and El-Diraby (2010a, p. 743) proposed an extendable construction process ontology, where “operations are modeled as interrelated processes or events that produce a product within a project, requiring resources with actor roles, mechanisms and constraints, attributes, and modalities” [133, 170]. Similarly, Benevolenskiy et al. (2012) developed a process configurator mechanism triggered by an ontology-based programming component and a rule-engine module to identify construction process patterns composed of resources, subprocesses, and building objects for process reusability from project to project [160].

While ontology-based studies contribute to better managing of construction processes by providing clear, structured, and unambiguous semantics while increasing process transparency among projects, they fall short in monitoring the real performance behavior of construction operations. In this regard, many taxonomies and ontologies in the construction domain for PMM are underutilized in practice, often because they try to represent dynamic process execution even though most of them are static in nature, meaning that it is time-consuming and labor-intensive to manually update them every time new knowledge process components need to be captured. An additional detailed analysis of top-down studies is provided in Table S5.

3.2.4.2. Bottom-Up PMM (Data-Driven) Studies

Sigalov and König (2017) presented an automated bottom-up technique to recognize typical construction process execution patterns as process templates. The proposed method uses a four-dimensional (4D) BIM model and a detailed project schedule as input. It then employs schedule decomposition algorithms and feature graph–based analysis to identify recurrent process patterns based on computed sub-schedule similarity indices to generate process templates [138]. Despite the automated method proposed in this study, the approach relies on an as-planned project schedule, as opposed to an as-happened process modeling approach.

Shi et al. (2008) introduced a task-based method that enables automated modeling of construction processes. The study focused on the standardization of tasks by creating a customized library of coded atomic task components that consisted of actions, methods, and objects, assembled into logical sequences to generate construction process models according to customer needs. Moreover, the process tasks were categorized into fully automated, semiautomated, and manual; they included all required parameters to be executed in a workflow management system (i.e., iFlow) [27]. Unlike other bottom-up studies, Shi et al. focused on modeling administrative processes, and although their logic could be applied to construction operations, they did not cover other BPM aspects such as automated process discovery, process evaluation, and process optimization, which are essential for process improvement [141]. Table 3-2 provides an aggregated comparative analysis of major bottom-up process-oriented studies in the AEC/FM domain, highlighting their main process of interest, focus, and main data sources, tools, and methods.

Table 3-2. Main PMM works – 'bottom-up' PMM approach.

References	Modeled Process	Data Source Details for Modeling Processes	Method or Instrument	Focus of the Study
Sigalov & König, 2017	Cast-in-place columns, Electrical and sanitary Installation	BIM and schedules (project information)	Feature-based similarity	Discovery & Reusability
Marengo et al., 2019	Floor and window installation	Historical project data	CoPModL (a modeling language for Construction Process Management)	Transparency & Conformance
Correa, 2018	Column assembly & masonry walls	Onsite sensors and historical data	Cyber-Physical System-Ultra Wide Band	Transparency
Amer & Golpavar-Fard, 2021	Various construction process sequence patterns	Schedules from previous projects	Text mining & supervised learning	Reusability
Liu et al., 2015	On-site construction of panelized building projects	BIM and project information; managers' experience	Discrete Event Simulation	Optimization
Shi et al., 2008	Procurement/ purchasing	Previous studies	Task-based programming	Reusability (Task Library)
Chau et al.2021	Contract Management & Scheduling - Highway construction	Historical project data	Pattern mining, statistical analysis, and network analysis	Pattern Discovery& Scheduling

Under the premise that existing standard process modeling notations fail to properly capture spatial locations where construction operations occur, Marengo et al. (2019) proposed a location-based process modeling language for the construction domain (CoPModL). CoPModL was built on formal semantics and a graph-based algorithm using linear temporal logic over finite traces theory to validate process model satisfiability. It considered elements like construction phasing, activities, task locations, and location-based task interdependencies. Marengo et al. (2019, 2021) offered an open web-based platform to model construction processes using CoPModL [80, 171]. Although CoPModL integrated the spatial location of process activities, the process modeling part was manual according to a pre-structured grid convention, which was error-prone and time-consuming. Table S6 provide a more detailed comparison of bottom-up studies.

3.2.4.3. Need for Automated Process Management

As a result of the top-down and bottom-up comparative analyses on process modeling and management, it turns out that construction projects in the AEC/FM domain have a clear need to transform current methods of capturing, modeling, and analyzing business processes from manual subjective techniques into more automated and data-driven techniques. Some studies have started to explore the capacity of computer vision and similar data-driven techniques such as machine learning and deep learning for activity recognition [172-175]; While such studies indirectly contribute to the automation of process discovery and analysis, they are set aside from the in-depth review analysis here, as this study specifically focuses on detecting, modeling, and management of E2E processes (unlike activity-oriented studies). For instance, Correa (2018) adopted an automated construction process modeling approach using Petri nets notation based on ultra-wide-band data acquisition as a CPS [176]. Amer and Golparvar-Fard (2021) automatically generated dynamic process model templates from construction schedules of previous projects through text mining and machine learning [177]. Similarly, Le et al. (2021) proposed a data-driven process modeling approach through pattern mining and network analysis to detect construction activity sequences based on as-built pairwise logical relationships in highway construction projects [140]. However, these studies focused merely on automatically modeling siloed construction activities without actively monitoring and analyzing all (i.e., E2E) process performance behavior to spot inefficiencies, which would have enabled a more holistic process monitoring and management approach.

Figure 3-6 reveals understudied research areas in the PMM literature for the AEC/FM domain, representing key opportunities for future research in this field. The coverage scale, defined based on a thorough subjective assessment, is as follows: (0 if the criteria are not investigated; 5 if the aspects are addressed by the study; and 1–4 if the study partially covers the subject). As seen in the figure, there is a need for longitudinal process-oriented studies that consider the entire BPM cycle, including the business processes redesign, the implementation of re-engineered processes, and, most important, the automated process monitoring with predictive capabilities to extract process performance behavior patterns so that they can be benchmarked against clearly defined business KPIs to trigger process improvement strategies and automation opportunities [56].

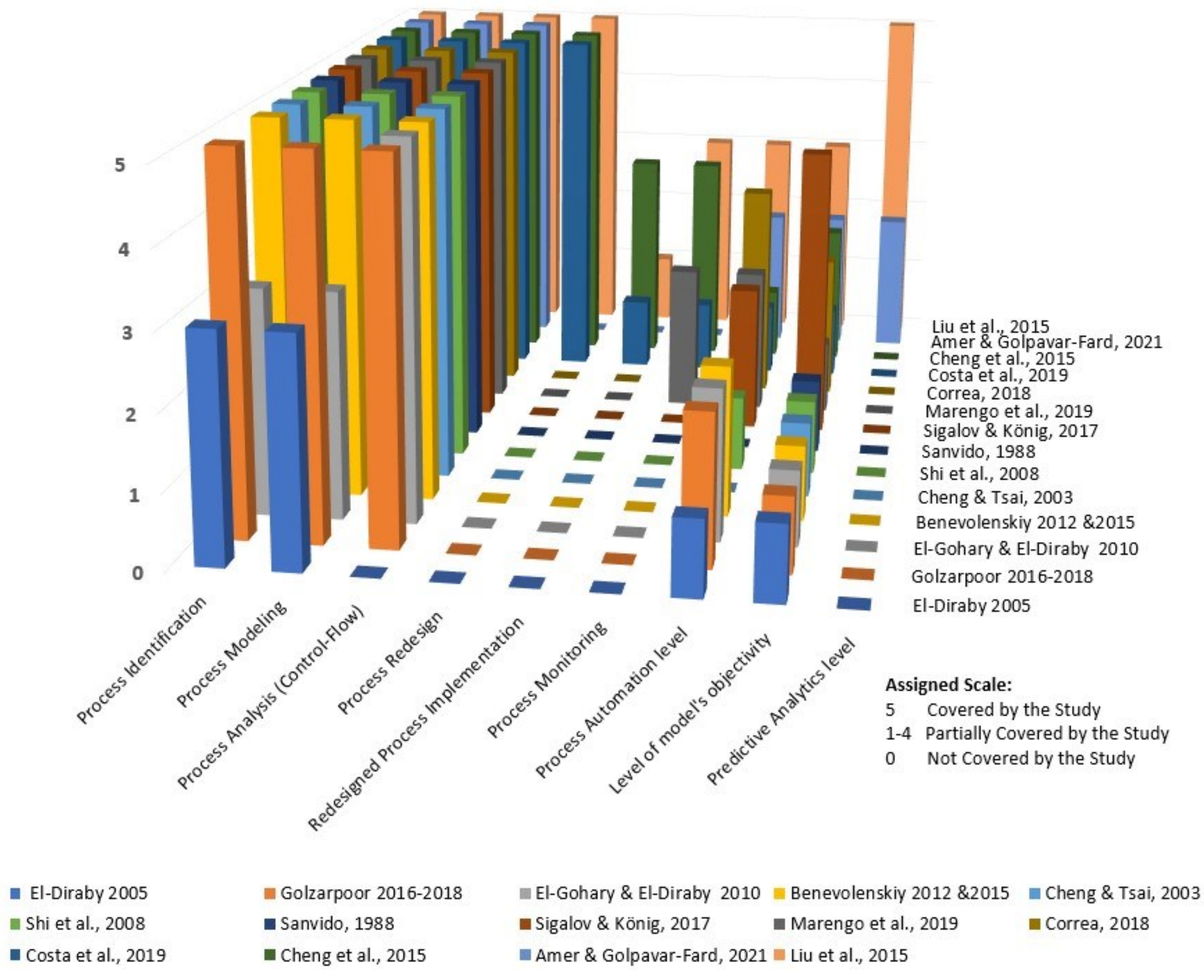


Figure 3-6. Aggregated perspective of PMM studies.

3.3. Process Mining in Construction

Taking the need for automation into consideration, process mining is all about logging, exploiting, and replaying actual operational data in the form of event logs on top of the automatically discovered process models to monitor and manage E2E processes [178]. Unlike digitally intensive sectors, construction is part of the lagging non-digitally intensive sectors facing stagnant productivity over the past three decades [3]. Hence, since construction projects are process-heavy, the adoption of process mining is becoming essential to address productivity stagnation in the sector, representing the missing piece for enabling more automated and efficient PMM in the AEC/FM domain.

Table 3-3 presents a comparative analysis of studies on process mining in the AEC/FM domain. These studies set the foundation for process mining in construction, representing early efforts to support automated data-driven PMM in the industry. However, the body of knowledge in this area is not yet mature. As shown in the table, these studies consider only one process perspective at the time of analyzing construction processes [89, 90, 178, 179]. This observation is important given that the recently emerged object-centric process mining discovery enables a multi-perspective cross-analysis of intertwined processes [180]. Thus, the feasibility of applying this technique should be further investigated. Also, the table shows that the construction-related event logs of these studies are majorly extracted from different data sources, such as BIM, unmanned aerial vehicle (UAV) technology, IT systems, facility management systems, and radiofrequency identification (RFID) systems, among others. It is important as well that the main process mining algorithms adopted by these studies to derive and analyze construction processes from operational event logs are the inductive miner and the fuzzy miner, which are among the most robust algorithms used by researchers. The former derives concurrency process behavior, while the latter facilitates discovery of a simplified sequential view of a process [11].

Table 3-3. Process mining studies in the AEC/FM domain by project phase

Study criteria\ Phase	Design		Construction			Operation
Author & Year	Kouhestani & Nik-Bakht (2020)	van Schaijk (2016)	van Schaijk (2016)	Pan & Zhang (2021)	Rashid & Louis (2020)	van Schaijk (2015)
Main Data sources	BIM (Model in Revit)	Systems Engineer IT system database	BIM & Construction Schedule As-Built 3D Point Clouds -UAV	BIM Models & Construction Schedule As-Built 3D Point Clouds -UAV	RFID Database	Planon Facility Management System
Applied Discovery Process Mining Algorithms	Inductive Miner	Fuzzy Miner	Fuzzy Miner	Inductive Miner Fuzzy Miner	Inductive Miner	Fuzzy Miner
Modeled Process Perspectives	Control-Flow Time Organizational	Control-Flow Time Organizational	Control-Flow Time Organizational	Organizational Time	Control-Flow Time	Control-Flow Time- Cost
	Single Perspective at a time	Single Perspective at a time	Single Perspective at a time	Single Perspective at a time	Single Perspective at a time	Single Perspective at a time
Modeled Process Disciplines	Architectural Structural MEP	Civil	Architectural and Structural	Architectural and Structural	Assembly Line 5 Units	Sanitary, Electric & Plumbing

Process Model Types	Design Authoring Process (As-Planned VS As-Happened)	As-Is Process Model (Design Specification Process)	As-Planned & As-Built (Construction Operational Process)	As-Planned & As-Built (Construction Operational Process)	Predefined Production Plan vs. As-Is Process (Assembly Line Process)	As-Is Process Maintenance Error Handling
Main Target Group	BIM Managers	Process Engineers Project Managers Contractors	BIM and Process Mining Researchers Industry Practitioners	Project Managers	Process Managers	Facility Management Processes
Main Enhancement Target	Design Productivity	Process Efficiency Process Transparency	Process Planning	Process Efficiency	Construction productivity, Process Improvement	Process Discovery, Process Improvement

Another important aspect to highlight in this comparative analysis is the fact that most of the process-oriented studies in the AEC/FM domain did not fully explore the potential of process mining technology across life-cycle phases. More specifically, the literature mainly covered process mining for design or construction operational processes [89, 90, 178] and barely addressed the potential of this technology for investigating the performance of key construction administrative processes such as requests for information (RFIs), procurement management, and change management. To tackle this gap, efforts have been made to exploit process mining automation in unveiling, analyzing, and improving the actual performance of construction administrative processes, particularly by developing a process mining use case for change order management during the construction phase and proposing a framework for event log generation specific to the construction domain [181]. Table S7 provides a detailed comparison analysis of process mining studies.

3.4. Identification of Construction Business Processes by Project Phase

The life cycle of most construction projects follows four major phases from conception to completion: tendering, preconstruction, construction, and postconstruction. In each of these phases, key administrative processes were identified from the literature review: process mining, modeling, and management as presented in Tables 1–3. I categorized these into three major classes: (1) project management–related; (2) financial management– related; and (3) sustainability–related. Accordingly, a conceptual classification process model, referred to as an IRIS model, was developed as shown in Figure 3-7. IRIS stands for integrated, reliable/reusable, intelligent, and sustainable, alluding to the process characteristics/capabilities that should be enabled and considered in any construction project.

Figure 3-7. IRIS classification model of service-based administrative processes.

More specifically, to generate this model, key administrative processes were categorized across a construction project's life cycle as reported in the literature, then validated and complemented them with subject matter expert inputs from the interviews. These processes include those in the tendering/planning phase, such as bidding and cost estimation [182-185]; the pre-construction phase, including VDC-based processes, procurement management, and risk management [186, 187]; the construction phase, comprising change order management, requests for information, and progress payments, among others [181, 188-190]; and the post-construction phase, such as deconstruction, commissioning, and work order management [31, 191]. In this context, the proposed model aims to build and identify a comprehensive process portfolio with the potential for implementing data-driven techniques supported by ICT (i.e., process mining) to facilitate performance monitoring, compliance assessment, and improvement initiatives by detecting automation opportunities for streamlining key business processes. It is worth mentioning that these processes may be interrelated with one another and across project phases.

3.5. Operational Model Canvas Design to Support PMM in Construction

Finally, a business process model canvas is proposed, derived from the exhaustive analysis of the literature on BPM, PMM, and process mining in the construction industry. Considering that construction processes tend to be overlooked across the life cycle of projects and the lack of visibility of real process performance, the proposed canvas shown in Figure 3-8 aims to address this issue by serving as a user-friendly and customizable tool for construction organizations and industry professionals to audit, monitor, and manage their construction operations. The canvas can help them build their corporate memory and information assets, including their main process portfolios, to better plan, document, assess, and communicate their key processes. It is also a supportive tool for identifying the need for business process re-engineering and performance monitoring (i.e., efficiency) of improved processes. The ultimate goal of this wireframe canvas model is to automatically create digital dynamic process performance dashboards considering multiple relevant perspectives and KPIs such as labor productivity, SNA, costs, and time process duration. It can support data-informed decision-making as an essential component for project and process controls, streamlined process certification audits [192], and automated PMM in the AEC/FM domain. The wireframe of the model presented here is composed of seven major components selected to be in line with the extracted high-level taxonomy depicted in Figure 3-4, the highlighted criteria from Figure 3-6, and informed by industry experts' inputs collected through semi-structured interviews, summarized in the next chapter (i.e., key business processes, major bottlenecks, and process data management needs). Notably, in the reviewed PMM literature no construction process-oriented digital dashboards were identified showing the components, whose monitoring is critical for carrying out process control audits and implementing longitudinal process improvement initiatives.

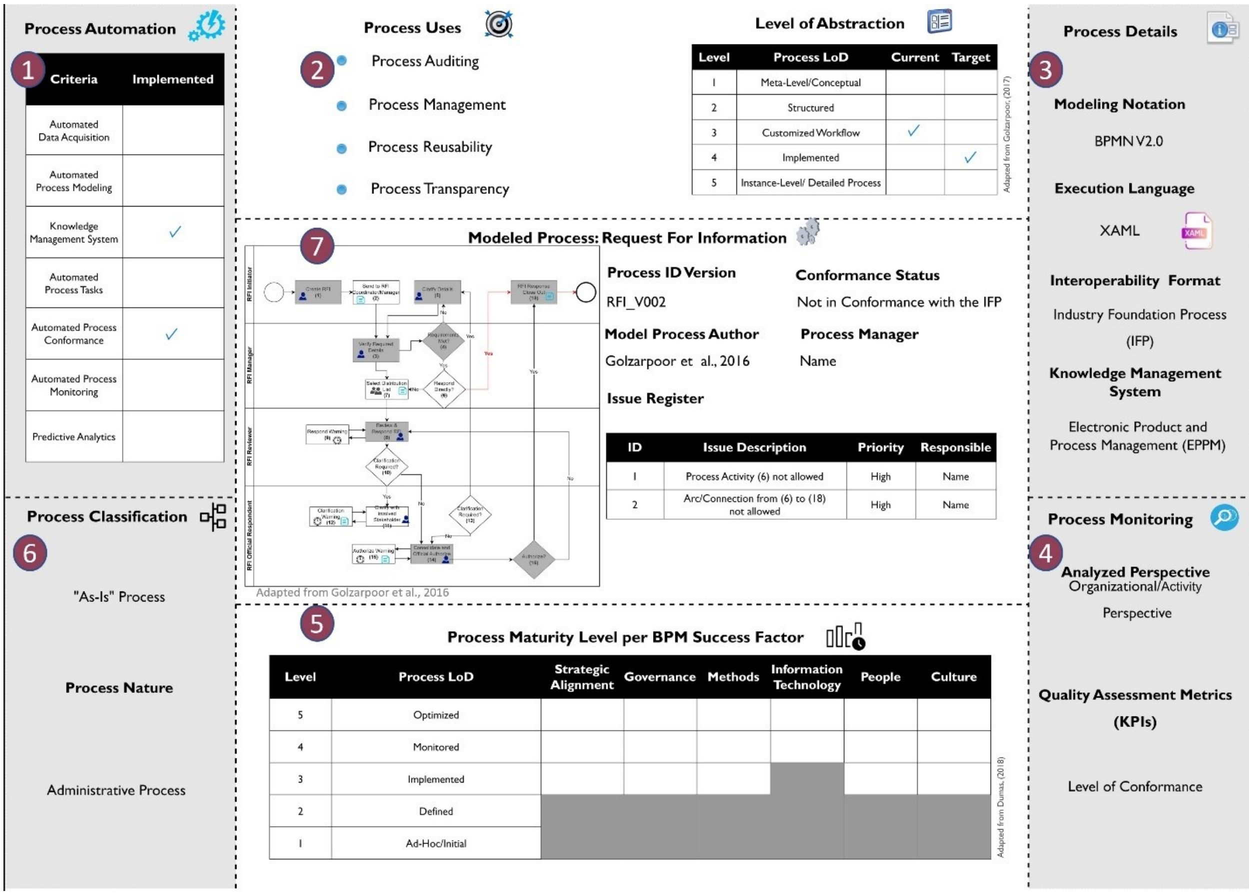


Figure 3-8. Business process model canvas management tool ("As-is" RFI process).

3.5.1. Process Automation (Section 1)

The purpose of this section is to monitor the current level of automation for the process of interest according to any automated aspects currently implemented to make modeled processes more efficient. These criteria were defined based on the reviewed literature, including but not limited to the automation of data acquisition for input process modeling; process discovery and modeling; repetitive or trivial process tasks; process conformance; process monitoring; implementation of knowledge management systems; and incorporation of predictive analytics. It is noteworthy that these criteria are customizable to meet organizational needs and goals.

3.5.2. Process Uses and Level of Abstraction (Section 2)

This section has two parts; the first part consists of specifying the intended uses of the modeled processes as shown in Figure 3-8. In parallel, the second part shows the current and target levels of abstraction considered for modeling the process of interest; these levels should be defined according to their intended process uses and organizational goals. In this vein, five main levels of abstraction were identified for process modeling suggested by [53]. From the most general to the most specific, they include the following: (1) meta-level: process steps are conceptually described either textually or in a flowchart; (2) structured level: process models are represented in a standard or formal modeling language notation including execution details; (3) customized level: structured processes are customized as required by construction organizations to ensure process quality and conformance with industry best practices, building standards, and organizational regulations; (4) implementation level: representation includes both manual and automated process tasks; and (5) instance level: representation includes the executed version of an implementation-level workflow, which can include several workflow instances carrying all the execution data associated with the process steps. Construction organizations must evaluate the current level of abstraction of their processes and define their target levels according to their operational requirements.

3.5.3. Process Information Details (Section 3)

At the top right of the canvas are the process modeling details, which can include formal modeling language notation, execution language, interoperable exchange file extensions, and any implemented management systems hosting digital traces of the processes under investigation.

3.5.4. Process Monitoring (Section 4)

At the bottom right of the canvas, two main aspects are considered: process model perspectives and, most important, key PPIs used to evaluate overall process performance. For the given example shown in Figure 3-8., the modeled process considers control flow and organizational perspectives; the key PPIs of interest are related to process conformance.

3.5.5. Process Maturity Level (Section 5)

This section of the canvas involves the evaluation of six success factors reported in the literature which form the basis of any BPM implementation initiative at different levels of process granularity. It should show the assessment of the maturity of the process of interest in terms of critical business aspects [52, 70], including how well the process supports the business strategy, meets regulations and incorporates efficient methods and relevant technologies, and how transparent it is for stakeholders so that it can be incorporated into a broader process-aware business culture.

3.5.6. Process Classification (Section 6)

Shown at the left of the canvas are modeled processes classified in terms of their depicted state and nature as described in previous sections of this study.

3.5.7. Process Model (Section 7)

This section represents the process model by its corresponding modeling notation and shows terms to describe the process, such as versioning, authorship, conformance status, auditor, and a register listing of all issues organized by priority and the person responsible for tracking and resolving them. It indicates what activities should be performed, by whom, and who can or cannot perform the process activities.

3.5.8. Canvas Implementation

To demonstrate the process canvas implementation, the RFI process benchmarked by Golzarpoor et al. (2016) is used as a reference. For this administrative process, the proposed model can, for instance, be used to monitor conformance checking. As shown in Figure 3-8, conformance checking is carried out by benchmarking the RFI “as-happened” against “as-planned” process model. The canvas helps to (1) identify non-conformities; (2) assign responsible parties to resolve conformance issues; and (3) set the priorities and business PPIs to track and monitor process execution flow. Moreover, it outlines (4) major process characteristics; (5) the analyzed perspectives; (6) the main uses of the documented process; and (7) a high-level assessment of process automation, abstraction, and maturity. Golzarpoor et al. (2016) extracted the as-happened RFI process model from previous process executions stored in a KM system, and automated process conformance checking was automated through an object-oriented programming alloy algorithm [24]. It is worth noting that the RFI process is present in most (if not all) construction projects and not managing it effectively can lead to work disruptions, additional cost, and time overruns. Figure S1 - Figure **S3** include a blank canvas and additional application examples of this process management support tool [26].

While process standardization is essential, no single model can fully capture the inherent variability of construction business processes, as reflected in the process taxonomy presented in Figure 3-4. Modeling, monitoring, managing, and ultimately improving these processes require consideration of specific situational factors, their dynamic nature, and the contextual attributes of the organization and its project portfolio. Without overgeneralizing, this operational process-oriented model provides the common denominator distilled from the literature, which must be tailored to specific project requirements and organizational management strategies when applied.

3.6. Identified Research Gaps and Future Directions

After a thorough examination of the PMM literature in the AEC/FM domain coupled with the results analysis of the industry-oriented interviews, the following major gaps and their potential future directions were identified:

1. The construction industry continues to rely on manual and subjective methods (i.e., workshops) for modeling and managing construction operations [75, 80]. Although these methods help to structure and model process workflows, the resulting models easily become outdated or they even end up on an office board not properly documented. This is time-consuming and error prone.

    - Investigate more advanced data acquisition technologies (e.g., IoT sensors, laser scanners, UAVs) and more novel automated methods (e.g., multidimensional process mining, AI-enhanced process mining, RPA) to derive, analyze, and improve construction processes from digital data sources (e.g., cloud platforms, BIM, KM systems). Implement them to streamline and objectively support construction PMM [76]. Identify and overcome barriers to adoption by analyzing and demonstrating, for instance, the real business value of implementing these processes as part of specific process improvement initiatives in construction organizations (return on investment in terms of productivity growth, cost, time saving, etc.).

    - Further explore process mining applicability in the digital construction supply chain for both traditional and modular construction to achieve significant process improvements [90, 179].

    - Realize the opportunity of leveraging process mining for financial business process operations (i.e., fraud detection) in the construction domain [181].

2. Most construction process models in simulation-based studies are as-planned rather than as-happen(ed)/actual, which can yield misleading simulation results.
   - Future construction simulation studies (e.g., DES, SD, ABS) can capture actual process execution performance (i.e., through process mining) to evaluate what-if process improvement scenarios that enable data-informed predictive analytics to optimize the actual behavior of construction processes.
   - Given the different nature of the two practices, develop separate applications for on-site and off-site construction. Process mining can be specifically helpful in aligning, monitoring, and improving off-site (factory) production processes.
3. There is a pressing need to operationalize sustainability governance strategies.
   - Investigate new policies, standards, and mandates for green procurement that have been defined and imposed by several developed countries to control and reduce embodied carbon emissions in the construction industry [189, 193]. Process mining plays a pivotal role in monitoring and ensuring compliance in the procurement decision-making process, with environmental regulations enabling green purchasing of goods and services.
4. Ontology-based studies do not usually capture the full E2E performance behavior of processes of interest, and it can be time-consuming to generate ontologies for them.
   - Consider adoption of a more hybrid approach so that the process semantics of existing ontologies (i.e., top-down) can be enriched with actual process performance behavior (i.e., bottom-up). A combination of these approaches appears to be the ultimate key to intelligent process modeling and management in construction. Research and development will be necessary to enable flexible process templates that can be fed by the operation data and self-improve based on analytics and/or human feedback. The proposed canvas can be extremely helpful in structuring data and information for achieving this goal.
5. Bottom-up PMM studies lack robust and quality process semantics as well as querying capabilities. Also, these studies are not longitudinal, meaning that they often do not consider

a holistic BPM management approach; instead, they are siloed process improvement efforts.

- Enrich bottom-up studies with more robust ontological semantics and with a broader scope of analysis to holistically support construction process management.

6. Existing process mining in the AEC/FM domain relies on a single perspective/case notion based on the XES standard instead of analyzing the performance behavior of intertwined business processes.

- Unveil and monitor the multidimensional behavior of business processes commonly found in construction projects, in accordance with recent international event log standards such as OCEL, thereby helping to untangle the inherent complexity of these processes.

7. Process mining studies in the AEC/FM domain analyze either design-based or operational processes, but they barely investigate the performance of significant service-based administrative processes, which can severely impact the timeline and cost of a project.

- Investigate process mining capabilities to derive and dynamically analyze the actual performance of service-based administrative processes across all life-cycle phases of a construction project.

8. The AEC/FM domain still lacks a robust and systematic methodological framework for generating high-quality event logs suitable for process mining applications.

- Gather and investigate industry process knowledge requirements and data needs to design a process-oriented framework tailored to the construction domain that can enable the automated generation of event logs and dynamic process performance analysis.

The common denominator in bridging these identified gaps is leveraging process mining capabilities alongside recent ICT developments (e.g., AI). One example is using process mining to automate process model discovery and re-engineering for the development of SOPs while continuously assessing and monitoring the performance of actual business operations to optimize their efficiency in the long run through data-driven decision-making and actionable insights.

3.7. Concluding Remarks

This study enriches the existing body of knowledge on process mining, modeling, and management in the AEC/FM domain by providing an exhaustive systematic review of three decades of research. The review includes an in-depth comparative analysis of process-oriented research highlighting the relevance that effective and efficient PMM has in achieving construction project success. In practice, construction processes (either administrative or operational) are frequently overlooked and the monitoring of their performance (if any) is performed by time-consuming manual means. Thus, if not properly managed and controlled, they can lead to productivity loss, financial stress, delays, or project failure. This work provides scholars and industry practitioners with detailed findings on PMM that highlight major research gaps, main drivers, and future directions.

The main contributions of this work can be summarized as follows:

- An exhaustive analysis of process modeling and management in the construction industry that expands the literature by offering a research classification framework with a holistic longitudinal view of construction business processes. This holistic view includes bottom-up and top-down approaches as well as major process modeling perspectives—as-planned, as-is/as-happened, and to-be”—for both operational and administrative processes. This framework will help future scholars better contextualize, benchmark, and drive PMM-oriented studies.
- The IRIS model introduced in Figure 3-7 contributes to the identification of key service-based administrative processes across major life-cycle phases of a construction project. The identification of these processes opens up new possibilities for studying and monitoring actual process performance considering automated means and methods. Also, it will help construction organizations build or enrich their corporate memories by having in place a portfolio of the most relevant construction processes and their key performance metrics to be monitored. This portfolio will serve as a basis for implementing future BPM initiatives to make processes more efficient and so ensure project success.
- A business process canvas model is proposed as a management support tool for project managers, owners, and project controllers in the AEC/FM domain to study, monitor, and

control the performance of the most critical business project operations. The aims are improvement, automation, standardization and repeatability, and interoperability of construction-related processes, ultimately driving construction productivity growth.

Despite these contributions, this study has the following limitations. One, the proposed conceptual business process canvas model was manually designed; future efforts should be made to automate the creation of digital process performance dashboards in real or near-real time. Two, although realistic examples of the applicability of the business process model canvas are provided (i.e., Figure 3-8), its practical application in construction projects remains limited, with several opportunity areas needing to evolve and be automatically generated in the near future. Three, Four, analysis of any empirical data to implement the proposed framework and the process model canvas remains out of the scope of this study. However, recent efforts have been made to automate the extraction of real process analytics from construction management platforms to monitor the performance of administrative processes and support data-driven decision-making across the project life cycle [181].

Given its limitations, it is noteworthy that the main goal of the synthesis offered here has been to bring the results closer to the application domain to enable process intelligence capabilities in construction projects.

CHAPTER 4. PROCESS KNOWLEDGE & OPERATIONS DATA CAPTURE[3]

As part of the requirements analysis (RO2), this chapter aims to investigate the significance of the identified administrative construction processes (i.e., service-based processes) by identifying those that are essential in most construction projects and exert a measurable impact on project time and cost performance. To this end, this chapter focuses on capturing process knowledge through the systematic collection of operational and contextual information related to the execution and significance of these processes across the entire construction project life cycle. The ultimate goal is to support the informed selection of critical process(es) for the detailed process mining analyses presented in Chapter 6. .

In line with the findings of the literature review and in response to the previously identified gaps (see Gap 4), the process knowledge collection strategy adopted a hybrid approach. This approach involved capturing and analyzing both tacit and explicit knowledge, as well as collecting and understanding real-world project operations data, given their combined relevance to the study. This strategy is illustrated in the high-level methodological flowchart presented in Figure 4-1.

4.1. Knowledge Capture

4.1.1. Semi-Structured Interviews with Subject Matter Experts

Explicit PMM knowledge drawn from the literature over the past three decades provides a robust

[3] This chapter is based upon:

A. J. Martinez Lagunas, S. Abbaspour, M. Nik-Bakht, and M. Ouf, "Digitalization and Process Management in Construction: A Current Industry Perspective," CMAA, 2025. https://www.cmaanet.org/sites/default/files/resource/Digitalization%20FINAL.pdf

Martinez, A., Nik-Bakht, M. (2024). "Enabling Process Mining in the Construction Industry: An Event Log Schema for Change Management Process." In: Desjardins, S., Poitras, G.J., Nik-Bakht, M. (eds) Proceedings of the Canadian Society for Civil Engineering Annual Conference 2023, Volume 3. *CSCE 2023*. Lecture Notes in Civil Engineering, vol 497. Springer, Cham. https://doi.org/10.1007/978-3-031-62170-3_8

Martinez Lagunas, A. J., & Nik-Bakht, M. (2024c). Process Mining, Modeling, and Management in Construction: A Critical Review of Three Decades of Research Coupled with a Current Industry Perspective. *Journal of Construction Engineering and Management*, *150*(11). https://doi.org/10.1061/jcemd4.coeng-14727 foundation for identifying core administrative construction processes that span major project phases, as visually aggregated in Figure 3-7. This also encompasses process components, intended model uses, modeling approaches, relevant PPIs, frequently reported inefficiencies, and other related elements. However, it is equally important to validate, align, and enrich this knowledge with a current industry perspective on PMM in the AEC/FM domain. This tacit knowledge is harder to capture as it mostly resides in the brains of industry practitioners [53]. Thus, it differs from person to person as it is highly dependent on the type of job position, project role, years of experience, and professional background. To capture this knowledge, semi-structured interviews were designed and conducted with subject matter experts.

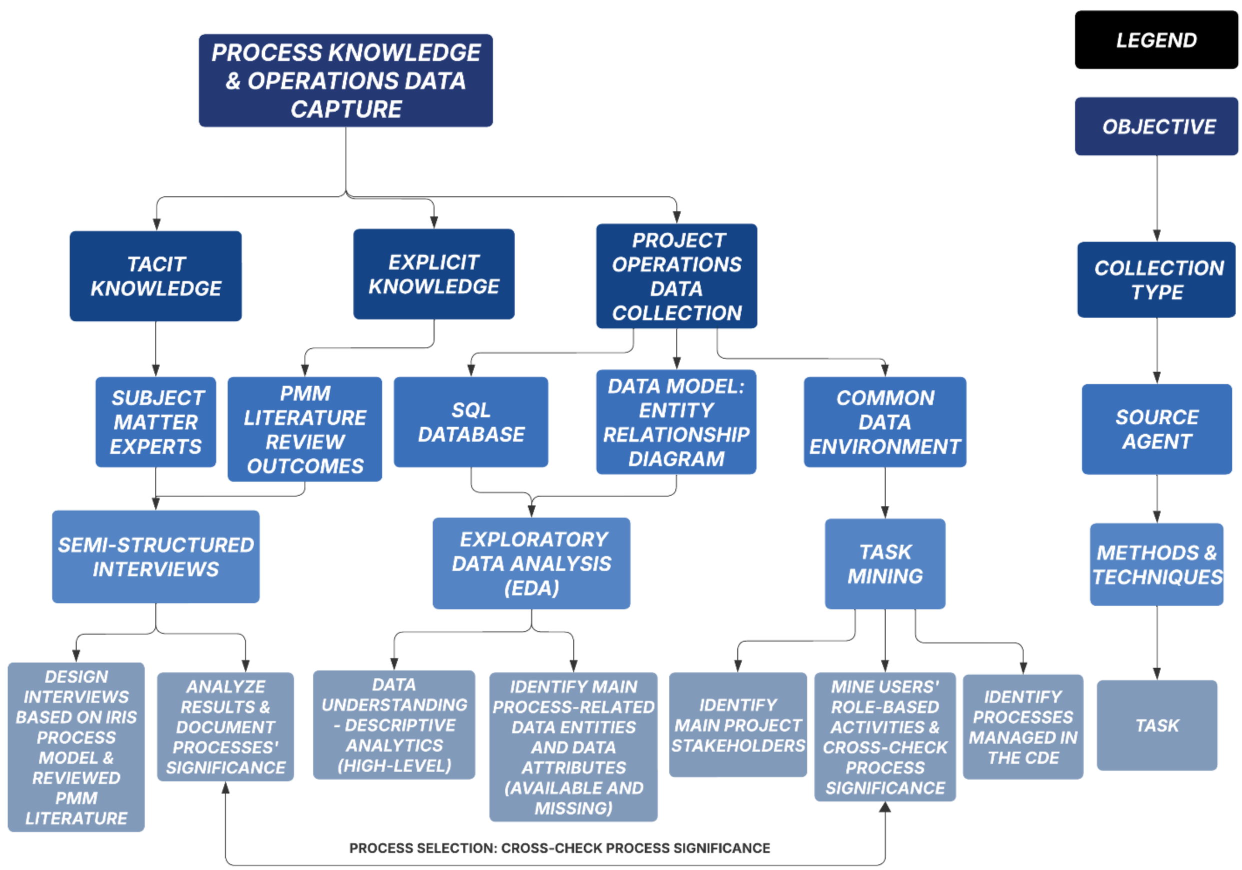


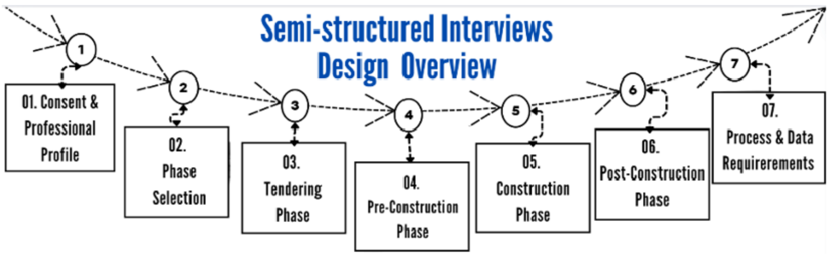
Figure 4-1. High-level methodology: process knowledge and operations data capture

4.1.1.1. Design Overview of Semi-Structured Interviews

The interviews were created, deployed, and stored using the Qualtrics online platform [194], under an ethics approval certificate by the Ethics Unit of the Office of Research, Concordia University

(Certificate number 30016847). The questions were formulated based on the exhaustively reviewed PMM literature and informed by the derived IRIS model in Figure 3-7. In total, 18 subject matter experts located in the US and Canada were interviewed. The interviewing period lasted three months, from mid-April to mid-July 2023. After careful examination, a maximum of 55 questions were included in the interviews to limit their duration to about an hour.

The interviews consisted of three major sections. The first section included questions about the respondent's professional background, questions at the project level, and questions at the company level to gain an overview of the number and type of projects managed by the companies the experts represented. The second section was the heart of the questionnaire as its questions were fully process-oriented to capture process relevancy and priorities across life-cycle phases of a construction project, main process workflows, required information exchanges, process KPIs, major types and sources of bottlenecks, and parties responsible for executing the identified business processes. The third section investigated industry operational needs or requirements of data-driven insights to support fact-based decision-making in construction projects. It is worth mentioning that the semi-structured interviews reported in this study were mostly conducted to verify and validate the completeness of the identified processes and modeling requirements. They aimed not to provide any prescriptive results but to collect process-related experience (i.e., tacit knowledge) from industry professionals on the basis of the framework extracted from the literature. A greedy method was used for approaching as many subject experts as possible. The criteria for selecting the respondents included (1) representatives from all project phases, (2) a wide range of years of experience in different key project roles and covering most of the main identified processes, and (3) experience working in several projects of different sizes and complexity. Following a semi-structured interview format, the respondents were given the flexibility to freely respond to several interview questions based on the results of the literature analysis, and to provide any additional inputs. Hence, although neither a predefined sample size nor a specific sampling technique was used, the responses were valid to highlight practical industry perspectives on PMM [195], on the basis of insights from the literature synthesis.



Section ID	Component ID	Content Focus On	Question(s) No.	Questions Focus On	Sample Interview Questions		
I	1	Consent & Professional Profile	1-11	Roles & Experience Involvement at Project Level Involvement at Company Level	**Q4.** What best describes your professional role? **Q6.** How many years of experience do you have in the construction industry?	**Q7.** What type of construction projects do you typically work on? **Q8.** What is the average base contract value of the construction projects you work on?	**Q10.** Approximately how many projects does your company complete per year?
	2	Phase Selection	12-15	Construction Project Phases	**Q12.** Our research has identified four major phases in a construction project which are: Tendering; Pre-construction; Construction; Post-construction. In your opinion, do these four phases represent adequately the lifecycle of a construction project? **Q15.** On which of these phases do you want to give more insights about?		
II	3	Tendering Phase	15-27	3 key identified processes	**Q16.** How would you rate the level of importance of the processes listed below with respect to your role? **Q24.** Which processes have the greatest impact on the project's completion timeline and requisites?	**Q21.** List up to 3 observed process bottlenecks or glitches? **Q26.** If applicable, what are the typical synonyms of each process?	**Q25.** What other phase(s) of the project would be highly affected by these processes? **Q27.** Which actors do you believe are associated with each process and what information is communicated between you and other actors?
	4	Pre-Construction Phase	28-46	10 key identified processes			
	5	Construction Phase	47-72	17 key identified processes			
	6	Post-Construction Phase	73-86	3 key identified processes			
III	7	Business Processes & Data Requirements	87-105	Industry needs and requirements on business process and data management	**Q87.** Has the volume of your project data increased over the last three years?	**Q90.** What percentage of your project data is available and suitable for decision-making? **Q99.** What are some challenges with using your data?	**Q102.** How much of your data is bad or unusable (ie:poor quality, inaccurate, untimely, incomplete or missing)? **Q103.** Why is the data poor quality?

Figure 4-2. Design overview of semi-structured interviews [46]

Furthermore, the interviews were designed in a way to cover key administrative construction processes across the project life cycle, as identified by the literature analysis. As shown in the interview design overview in Figure 4-2, questions in Section Components 1 and 2 covered the respondents' professional profile and project life-cycle phases [196]; Section Components 3 through 6, identifying the major processes in the different project phases, were informed by the reviewed literature. For instance, Component 3 comprised questions related to processes in the tendering phase, including cost estimation [184], bidding [185], and budgeting [182, 183]. Similarly, Section Components 4 through 6 comprised questions related to key processes reported in research studies in the pre-construction, construction, and post-construction phases. Questions were designed based on analysis of the studies provided in Table 3-1-Table **3-3** and S1–S9. Questions were formulated to helped to gain insights into aspects such as officially modeled processes [36, 53, 184]; process model categories, i.e., “as-planned”, ”as-is”, “as-happened”, or “to-be”, taken from [93, 128, 148] and identified process inefficiencies [53, 197]. Examples of these questions included (1) rate the importance of those processes in each phase; (2) describe the ideal (i.e., as-planned) process execution and its core steps; (3) identify main information requirements and stakeholders in key construction processes; (4) list main observed process bottlenecks or pain points; and (5) estimate potential project impacts due to inefficient construction operations among others.

Outcomes of the literature analysis were used to structure and articulate questions for the interviews in different ways. First and foremost, the main business processes in construction were aggregated from the literature analysis and classified as service-based administrative or product-based operational. Additionally a list of subprocesses and workflows was created for various phases of a project’s life cycle based on evidence from the literature [184, 196]. The list was then completed based on the information collected during the interviews. Second, for each class two comprehensive lists were developed: (1) shortcomings, challenges, and barriers listed for the processes in the literature; and (2) opportunities and solutions offered in the literature for process modeling, management, and improvement. Next, given the intrinsic intertwining of construction processes across a project’s life cycle [45], questions regarding the ripple effect of processes and their intertwinement across project phases were included. Then the important aspects emphasized in the literature to improve PMM in construction were translated into questions. One of the most frequently emphasized needs in the literature was the adoption of digitalization and development of data-driven solutions to support PMM [160, 178, 197]. Accordingly, another section component to the interview was added, Component 7, to ask respondents about the digital maturity of organizations in business processes management, mandates (i.e., SOPs), and business needs. Finally, a few questions were added regarding the most commonly used lexicon across the industry related to the identified business processes for future taxonomy and ontology development [133]. Interview outcomes are summarized and described next.

4.1.1.2. Summary of Interview Results

Based on the derived IRIS model introduced in Figure 3-7, subject matter experts were interviewed regarding their professional profiles and project-related experience. They were then asked to select a construction project phase aligned with their roles and expertise. Similar questions were formulated for each identified process in each phase. The main process-related questions included relative process importance (Q1); potential impact of inefficient processes on the project timeline (Q2); and identified process bottlenecks (Q3).

The literature and the initial data collected from industry experts suggested that the life cycle of construction projects follows four major phases once project feasibility has been verified: tendering, pre-construction, construction, and post-construction. Normally, tendering and post-construction are less process-intensive than pre-construction and construction. Thus, a balanced number of industry experts were interviewed proportional to the number of identified standard processes in each phase, as shown in Figure 4-3. More specifically, 18 industry professionals were interviewed, including 3 in tendering, 3 in post-construction, 6 in preconstruction, and 6 in construction. This section presents the main results of the interviews. The first part discusses the respondents' professional profiles. Figure 4-3 shows the years of experience of each respondent per role and phase. It is important that some of the respondents have had multiple roles throughout their professional careers. At the project level, Figure 4-4 depicts the main project types on which respondents have worked against average contract values.

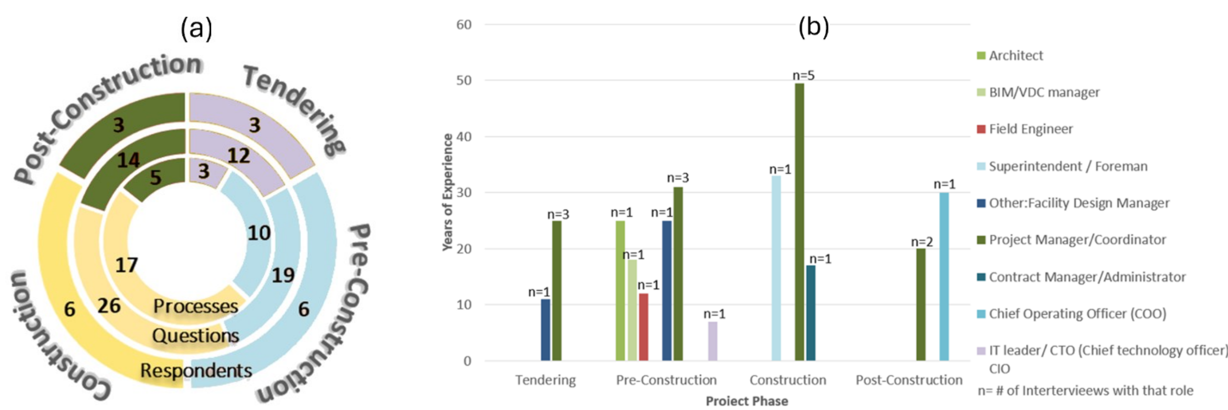


Figure 4-3. Interview composition and professional profiles: (a) structure and distribution; and (b) years of experience by role.

Figure 4-4. Project types (response counts) by project base contract.

At the company level, Figure 4-5 provides details about the number of projects managed per year by the companies the respondents work on, highlighting the average dollar value of the projects and the project life-cycle phase in which the respondents have been involved.

		Project Base Contract Value						
Project Phase	Max. No.of Projects	$100K - $500K	$500K - $1M	$1M - $5M	$5M - $20M	$50M - $100M	$100M - $1B	$1B and over
Tendering	10-25			33.33%				
	25-100	33.33%						
	100-1000							33.33%
Pre-Construction	Less than 5					40.00%	20.00%	
	5-10				20.00%			
Construction	Less than 5			16.67%	50.00%			
	100-1000						16.67%	16.67%
Post-Construction	10-25			50.00%				
	100-1000		50.00%					

% of Respondents 16.67% – 50.00%

Figure 4-5. Respondent support (%) by project contract value, phase, and number of managed projects per company per year.

Process Relevance and Bottlenecks – The second part of the interviews was fully process-oriented, Figure 4-6 shows the most relevant administrative processes according to respondents in terms of relative importance and potential impact on project timelines. Percentages in parenthesis indicate respondents' support. These outcomes help to identify critical priority construction business operations from the practical industry perspective, where process performance requires frequent monitoring using automated techniques (i.e., for future process mining use case development). As observed in the figure, there are important relations and variations between the significance of these processes and their potential impact on a construction project's completion timeline. In the case of tendering, the interviewees unanimously emphasized the "Bidding" process is the most preeminent business operation with 66% of the respondents stating that it can have an important effect on the project's completion timeline. Given that the foundations of project planning are laid during the bid preparation, the synergy is not surprising. Under the pre-construction phase, four out of ten identified processes stood out in terms of significance, i.e., "**Contract Management**", "Design Authoring/Design Approval", "Procurement Management", and "Submittal & Transmittal Approval". Several factors may contribute to this, yet a salient evaluation from a time approval perspective suggests that the Submittal and Transmittal **approvals** often necessitate frequent stakeholder involvement and decision-making to facilitate the seamless process progression (83%). This is closely followed by the importance of design approvals and procurement management, while **issues related to contract management may have critical implications on the project completion timeline**.

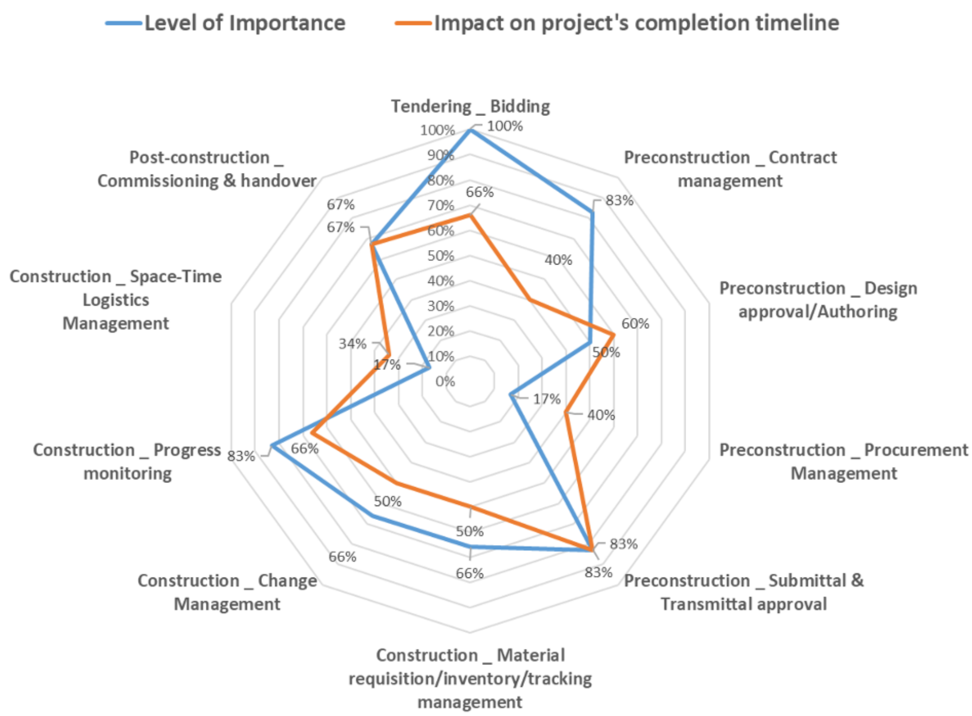


Figure 4-6. Process relevance and their effects on projects' completion timeline according to respondents

Under the construction phase, the three most important business operations according to respondents (out of a total of 17 processes) include "Progress Monitoring", "Material Management", and "**Change Order Management**", which can also have a **high impact on the overall timeline** of construction projects as they primarily deal with 'day-to-day' or 'week-to-week' actions and approvals aimed at controlling project progress and ensuring adherence to the schedule. At a lower level of respondents' support, "space-time planning" of construction operations is recognized as another critical process during the construction phase. Last but not least, at the end of the construction phase and during the post-construction phase, the "Commissioning & Handover" process holds a significant mid-level position in both regards. equally rated at 67%. It shows that stakeholders may perceive a balanced risk for this process, which can be more manageable.

To understand the operational challenges and execution constraints associated with these processes, the main process bottlenecks perceived by the interviewed industry experts for high-impact processes of each phase are presented in Table 4-1. These bottlenecks can have a direct impact on project timelines that in turn result in cost overruns. For instance, inaccurate estimations, overpricing, and inadequate material tracking during the 'Bidding' and 'Material Management' business operations diminish process efficiency, increase costs, and potentially lead to significant project delays. These bottlenecks shed light into the root causes of poor process performance. However, it is important to verify these bottlenecks or to unveil some others that were not identified through more automated and data-driven methods.

Table 4-1. Observed identified process bottlenecks by respondents per project phase

Phase	Process	Q3 Identified Bottlenecks?
Tendering	Bidding	Missing or Incomplete Drawings; Contractors Overpricing; Unreliable and inaccurate Estimations
Pre-Construction	Contract Management	Legal department revision; information management; BIM not specified in the contract; Inexperienced Project Manager
	Design Approval	Poor communication; lack visibility on approvers; delays in revision times; delays in approval times
	Submittal/ Transmittal Approval	Poor track of logs; manual workflows; revision delays
	Procurement Management	Lack of monitoring of vendors' prices; poor visibility on vendor agreements; foremen not ordering from partner vendors
Construction	Labor/Resource Management	Lack of visibility of who is doing what, where, and when
	Material Management	Unplanned daily orders by supervisors; poor inventory management; ordering much more than needed; poor material tracking
	Construction Progress Monitoring	Reporting more progress than the real one; poor visibility on remaining work and time to complete it; Overlapping activities too much
	Site-Space Logistics Management	At peak time conflicting cranes/machines

	Change Management	Getting all answers at the right time from different parties involved; poor coordination; Contract clauses not specific enough, scope and exemptions not detailed enough; Getting clients' approvals; change revisions; delays in estimating and submitting change
Post-Construction	Commissioning & Handover	Lack of process visibility; Non-standard punch list process; Inexperienced Owners

Concerning the identified processes outlined in Figure 4-6 and Table S8, the interviewees were also tasked with identifying the main process stakeholders' roles linked to each process and delineating the nature of information exchanged during the execution of these processes as well as the most common used lexicon in the construction domain to describe these processes, as detailed in Table 4-2 and Table S9, respectively.

Table 4-2. Non-exhaustive list of process stakeholders and information exchange.

Phase	Most important process	Process Stakeholders	Information exchange/ communicated
Tendering	Bidding	bidders, professionals	posting of tender, bids
		contractors and subcontractors	consultant report, plans, tender documents, specifications, quotes, schedules, shop drawings
		legal department, risk managers	insurance, liabilities
Preconstruction	Contract management	Project manager, Business unit leader, client	deviations from standards or initial tendering language, uncommon languages or payment terms
		owner, designer	contract, scope of work
		PM	BIM annexure
		developer team	scope of project
		PM, Finance	base contract value, value of changes
	Submittal & Transmittal approval	designer, contractor	construction docs
		PM, PD, BIM lead	BIM in compliance to client standards
		PM, design team, leaders (directors)	documents that need to be submitted, packages, documents
	Construction progress monitoring	site team, owners (client)	updated schedules, delays, advances, hurdles etc.
		surveillant	progress data
		foreman	forecasting details, FTC or FAC or %

Construction	Change Order Management	Owners, consultants, GC, speciality contractors, PM, PC	It can involve change events, change orders, prime contract change orders, Site Instructions (SI), contemplated change notices (CCN), change requests (CR)
Post-construction	Commissioning & handover	Landlord/owner, GC, facility managers	deficiencies list, premises' review, lease amendments
		service provider (commissioning agent), Business product manufacturers	requirement checklist, as-built drawings, reports for major systems,
		PM, project coordinator, suppliers, subcontractors, engineers, shop manager	deadline, eng. schedules SB/SC, maintenance doc, warranty doc

Cross-Phase Process Intertwinement – An aggregated view of the identified business processes across each project phase is shown in Figure 4-7. These processes were identified based on an exhaustive analysis of the PMM literature, validated by industry experts. The interviewees were tasked with indicating the process interplay, i.e. the degree to which each process contributes to affecting other phases and other processes. This diagram was derived from analyzing the interviewees` responses to questions such as those described in Figure 4-2. The three main parts of the diagram show the distribution of processes across different project phases, the number of responses supporting the importance of each process, and the percentage of cross-phase interplay. The connections demonstrate the flow and influence of each process from one project phase to another, indicating which phases are most affected by these processes and to what extent. The cross-phase analysis presented in Figure 4-7 reveals that with 32% of all identified processes, the pre-construction phase bears the highest number of process interplay indicating that decisions or actions taken in pre-construction are highly influential on the later construction activities.’ support.

As noted in Figure 4-7, the construction phase is referred by respondents to hold the 2nd highest cross-phase process interplay with 25.6 %. the “Procurement management” process during this phase has interrelations with the tendering phase (43%), pre-construction phase (43%), and post-construction phase (14%). Moreover, it is important to note that the “Change Order Management” process (during the construction phase) is strongly intertwined with the preconstruction phase, with a support share of 75% for that specific process, as it constitutes a subprocess within the broader “Contract Management” process. Similarly, the "Material

Management" process plays a significant role in the project`s success. Late material deliveries or untimely requisitions can directly impact the project timeline. Therefore, effective material management during construction also results from careful financial and resource planning during pre-construction. Besides, the "Material Management", "Procurement", and "Request for Information (RFI)" processes are interrelated not only with one another but also, they can be influenced by other processes such as "Design Approval/Authoring", "Submittal/Transmittal Approvals" and “Change Order Management”. For example, a delay in the approval of a material-related submittal can influence both the procurement of that material and the response time of an RFI, potentially leading to construction work disruptions and costly changes.

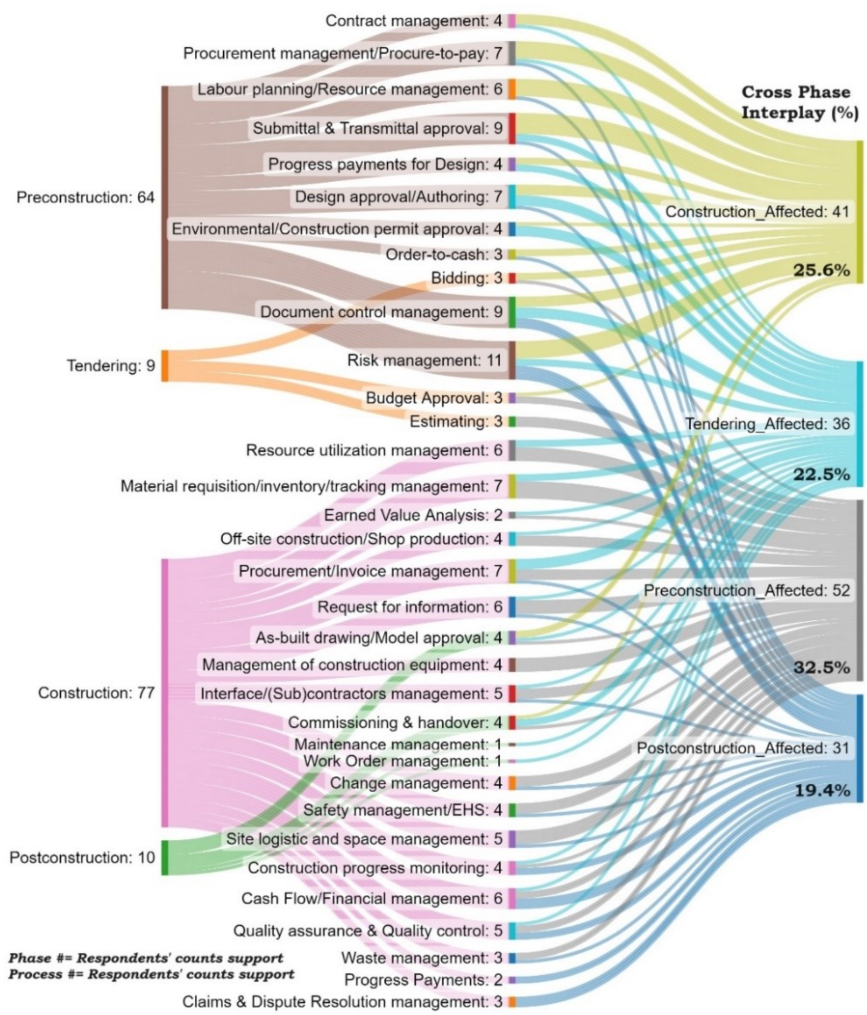


Figure 4-7. Cross-phase process impact and interplay according to respondents.

The results of the interviews provided the practical perspective of industry practitioners to validate the theoretical findings from the PMM literature analysis, providing critical insights about (1) significant service-based administrative processes broken down by project life-cycle phase; (2) criticality of identified processes in terms of management importance and potential impact on their completion timeline; (3) major observed process inefficiencies/bottlenecks; (4) main parties involved in the execution of these processes; and (5) common workflows and types of information exchange for each process. This tacit knowledge set the foundation for future process mining use case development related to the most significant construction processes so that the study findings can be aligned against industry needs and requirements.

As presented in Figure 4-6 and Figure 4-7 , one key finding from the interviews was the validation of the in-depth literature analysis regarding the identification of the most important service-based administrative construction processes across project lifecycle phases, along with their cross-phase and cross-process interrelations, as informed by experts' knowledge and practical insights. As illustrated in these figures, effective contract management plays a critical role across all phases of a construction project, directly contributing to its successful execution and delivery.

During the planning and tendering phase, contract terms are typically specified in key tender documents such as the general conditions, special provisions, and draft agreement forms. In the pre-construction phase, contract administration becomes essential for managing design changes, addressing project risks, and coordinating procurement activities.

Throughout the construction phase, contract management is closely linked to Construction Change Order Management. Change orders refer to formal contractual modifications to the original agreed scope of work[198]. The effective management of these changes is essential to ensuring timely and compliant project execution whenever adjustments are required. In the post-construction phase, contract management ensures proper project close-out, with commissioning and handover processes playing a vital role in meeting and fulfilling contractual obligations, including technical specifications.

On the other hand, information exchange and project communication management processes such as submittals, transmittals, and RFI management are also critical. These processes support streamlined stakeholder collaboration and help ensure seamless project execution. They also contribute to reducing construction changes by defining clear project briefs, specifying information requirements, and clarifying scope early in the project.

As highlighted in Figure 4-7, the Change Order Management process is strongly influenced by information management processes due to their close interrelationship. When poorly managed, processes such as RFI management can lead to misunderstandings or gaps in information. These issues often trigger construction changes that negatively affect overall productivity, delay the associated construction work and corresponding contractual progress payments, and may ultimately result in significant project time and cost overruns.

Despite these benefits, the semi-structured interviews did not rely on a strict sample-selection procedure. Instead, a pragmatic participant recruitment approach was adopted, prioritizing diversity in project roles and breadth of experience across multiple project phases and processes. As a result, the validation is limited to the 18 industry representatives interviewed; expanding the panel in future work could further enhance the robustness and depth of the insights. For instance, some of the respondents admitted to not having enough experience and qualifications in financial processes. Thus, to draw prescriptive conclusions, future studies should follow a systematic survey design and execution and must consider expanding the interviews to include subject matter experts related to those processes. Nevertheless, it should be noted that the interviewed industry representatives recognized the high potential of data-driven decision-making in enhancing process efficiency and facilitating cross-phase process management, enabled by an effective corporate data management strategy that ensures high-quality and usable project operations data ready to support these decisions.

4.2. Project Operations Data Collection

In addition to the tacit process-oriented knowledge captured through semi-structured interviews with subject-matter experts and the explicit knowledge derived from an in-depth review of the PMM literature, project operations data was collected from out industry partner, as illustrated in Figure 4-1. The primary objectives of this data collection were threefold: (i) to establish a solid foundation for the design and application of the proposed process-oriented method presented in Figure 5-1 ; (ii) to cross-validate, using a bottom-up approach, the interview findings regarding the relevance and applicability of the Change Order Management process and its interrelated subprocesses, as described in Section 4.1.1.2; and (iii) to assess data completeness and suitability for the implementation phase of the study. This latter objective included identifying primary data agents (i.e., core source systems), detecting critical data gaps through verification of the availability of key data attributes and their corresponding values required for E2E process-oriented analysis, and developing a robust understanding of the underlying data structures and relationships supporting the selected process.

It should be noted that, although the project operations data discussed in this chapter differs from the dataset analyzed in the case study presented in Chapter 6. , both pertain to interrelated processes. This chapter focuses on developing a high-level understanding of the underlying data structures supporting the Change Order Management process, whereas Chapter 6. concentrates on the quantitative assessment of the Request for Information (RFI) process, which constitutes a closely linked and operationally interdependent subprocess, using the methodology proposed in Chapter 5. At this early stage, the primary objective of this chapter was to gain a deeper understanding of the two collected initial data sources by conducting exploratory data analysis (EDA) and systematically examining each data entity, its attributes, and their interrelationships from a process-oriented perspective. This data-understanding phase laid the foundation for the subsequent development of process mining use cases and case study investigations, as most medium- to large-sized construction contractors in North America manage their project operations data through similarly structured, data-oriented architectures (i.e., SQL databases populated with data originating from various information systems).

Initially, this collection comprised two primary data sources: (i) a SQL database with raw project operations data managed by a large North American specialty contractor (SC), with this data being pulled from multiple source digital systems; and (ii) a data model (DM) that underpins the structure, relationships, and stored procedures governing the SQL data, serving as the blueprint for how data attributes, their data types, and their values are organized and interrelated to support operations.

4.2.1. Foundational Data Model: Entity Relationship Diagram (ERD)

The underlying data model (DM) for the construction project operations being managed by the large SC is represented in the ERD shown in Figure 4-8. Specific data entities related to the

Change Order Management process were identified in this diagram, along with their existing data attributes, data types (e.g., categorical, numerical), linking attributes, and cardinality relationships. A construction change order refers to any formal modification (addition or subtraction) to the original project contractual scope of work [35]. Accordingly, the associated data entities in the ERD include projects, contracts, change events, change orders, workers, project roles, and tracked working time. With respect to the data entity relationships and their cardinality, as shown in Figure 4-8, a construction project can be associated with one or multiple contracts. A contract can include multiple change orders, and each change order may contain several change event items. Similarly, various project stakeholders, each assuming one or more roles, may be involved in multiple change orders, and the time tracked in executing these changes is recorded in the tracked time entity, which serves as one of the inputs for managing subsequent progress payment applications/requests. Further details on the existing data attributes from these entity tables are provided in Section 4.2.2.

Figure 4-8. Excerpt of the Entity Relationship Diagram (ERD) - Change Order Management Process

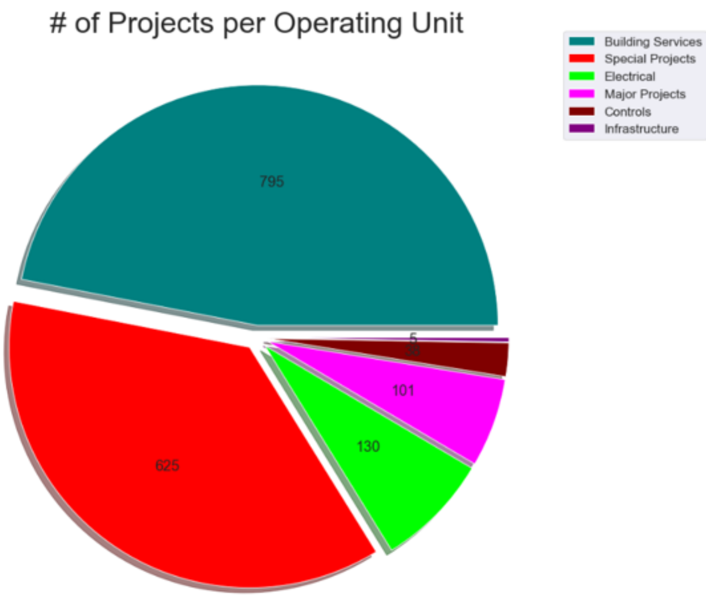
4.2.2. High-level Exploratory Data Analysis (EDA) of SQL Project Data

Initially, the static snapshot of the SQL database managed by the large specialty contractor contained data related to 3,439 MEP (Mechanical, Electrical, and Plumbing) construction projects and more than 88,000 change orders. The SQL database stored procedures were structured based on the ERD presented in Figure 4-8, which specifies the main attributes, associated data types, and the key entity relationships. An excerpt of this raw SQL data is summarized in Table 4-3, indicating the number of data records contained in each table, the number of attributes they included, their main unique identifier attributes (primary keys), and the linking relational attributes between data tables (foreign keys). For this specific large MEP contractor, most data tables were retrieved from Procore [199] with the exception of the Contracts table that was synched to Procore but initially pulled from Oracle NetSuite [200] and the Tracked Time table coming from both Procore Field Productivity [201] and Rhumbix software platforms [202].

Table 4-3. Summary of raw SQL project operations' data

Data Table Name	Datapoints	Attributes	Primary Key	Foreign Keys
Projects	3,439	37	ProjectInstanceId	ImportId; ProjectId
Project Roles	8,180	18	ProjectRoleInstanceId	
Contracts	4,089	25	ContractInstanceId	
Change Events	25,737	23	ChangeEventsInstanceId	
Change Orders	88,097	23	ChangeOrdersInstanceId	
Tracked Time	4,099,703	34	TrackedTimeInstanceId	ImportId; ProjectId; WorkerId
Workers	8,716	36	WorkerInstanceId	WorkerId; ImportId
Paid Time	914,845	26	PaidTimeInstanceId	ImportId; ProjectId; WorkerId

Among these projects, 1,694 were actively managed by the SC's construction operations department and had their primary operating unit explicitly identified, as shown in Figure 4-9. In addition, according to the SC's project operating unit classification, special projects referred to retrofit work whose schedule durations typically were less than one year and their project values below 1 million CAD. Major projects included new construction work with longer contractual durations and higher contract values, while large infrastructure projects primarily comprised new MEP construction for healthcare facilities, with project values of several hundred million CAD. The Building Services category was reserved for bundled, contracted MEP service activities required during the post-construction phase, while Electrical and Controls projects involved discipline-based, narrowly defined scopes of work.



Filtering out Projects with very low/ zero Project Base Contract Values

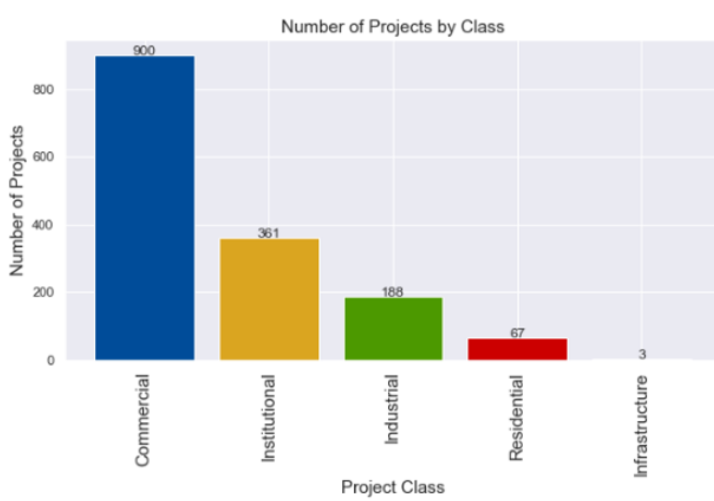
	ProjectOperatingUnit
Building Services	795
Special Projects	625
Electrical	130
Major Projects	101
Controls	38
Infrastructure	5

Are part of

	ProjectDepartment
Operations	1694

Figure 4-9. MEP Construction Projects by Operating Unit

Moreover, as shown in Figure 4-10, five project types were identified in the analyzed dataset, namely commercial, institutional, industrial, residential, and infrastructure, where such classification information was available.



	ProjectClassification
Commercial	900
Institutional	361
Industrial	188
Residential	67
Infrastructure	3

Figure 4-10. Construction Project Classification

To provide an overview of project classification by base contract value, projects were further categorized into low- and high-value segments based on the statistical quartiles of their base contract amounts, as shown in Figure 4-11 and Figure **4-12**. The low-value projects observed in Figure 4-11 were largely concentrated below approximately 1 million CAD and were predominantly classified as special projects, with commercial and institutional projects exhibiting lower median base contract values and a higher frequency of observations. In contrast, Figure 4-12 shows that high-value projects generally ranged from approximately CAD three million to over 20 million CAD, with institutional, industrial, and residential project types exhibiting higher median base contract values and greater variability.

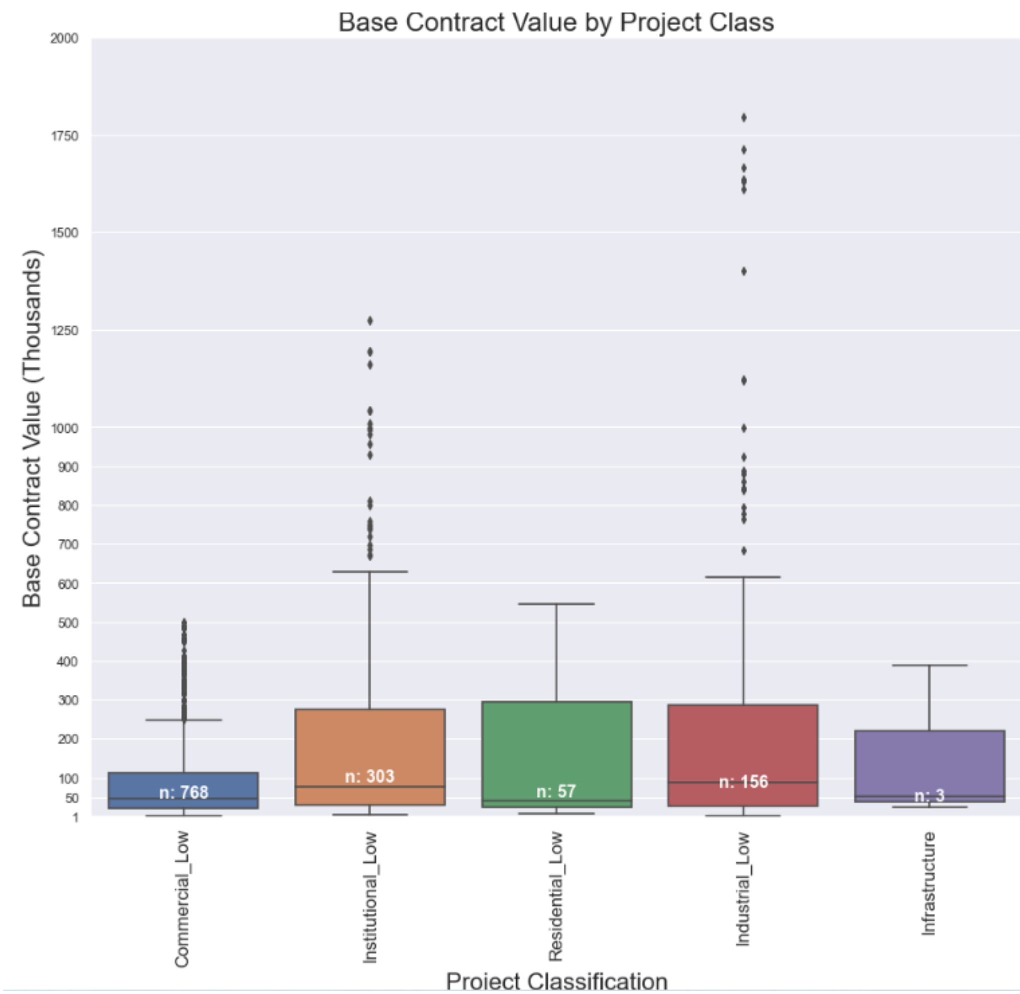


Figure 4-11. Projects by Base Contract Value Segments: Low Segment

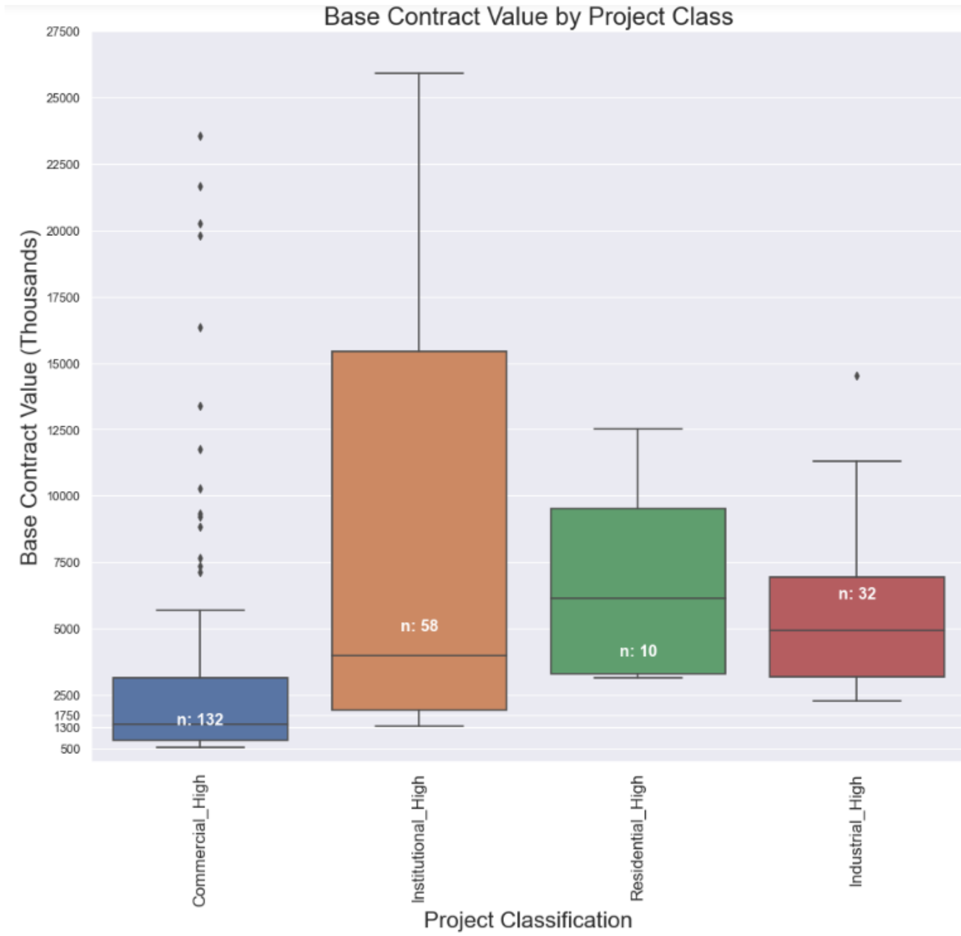


Figure 4-12. Projects by Base Contract Value Segments: High Segment

It should be reiterated that the goal at this point was primarily to understand the type of project operations data available in the SC's SQL database associated with the Change Order Management process, as well as to verify its completeness for subsequent process mining–based analysis presented in the following chapters of this thesis. With this in mind, main stakeholder roles and the corresponding number of workers per role involved in tracked MEP construction project operations, including those participating in the construction change order management process, were identified based on the project roles, workers, and tracked-time tables. For instance, for a major project, these roles included Project Managers (PMs), Project Coordinators (PCs), MEP Engineers, Project Administrators, Executives, Owners, Financial Managers, General Foremen, Superintendents, and various specialty trade worker roles, among others , as shown in Figure 4-13.

Figure 4-13. Excerpt of Number of Workers by Project Roles

From the 1,519 projects with non-empty or non-low contract values, approximately 19,100 change events and 43,980 change orders were identified in the corresponding data entities, from which the main reasons triggering these change events, as well as the associated change types, were identified as shown in Figure 4-14. 'Site Instructions', 'Design Development', 'Change Notices', and 'Client Requests' were identified as primary drivers of project changes. 'Site Instructions' refer to formal written directives, typically issued by construction managers, engineers, technical consultants, or architects, to clarify, modify, or direct the execution of work [47]. In addition, the primary predefined change types included 'Owner Changes', 'Change Directives', and 'Contingencies'. Such classifications are commonly adopted by AEC/FM organizations to organize issued changes according to budget cost codes. A change directive is typically issued by the project owner or an authorized representative when agreement on the cost of a change has not yet been reached with the general contractor or subcontractors. It authorizes the immediate implementation of field work associated with disputed changes [198]. Contingencies are normally included in the original contract amount to account for unknown items that may result in additional costs not specified in the contract[19]. A glossary of similar terms is provided in Appendix S10.

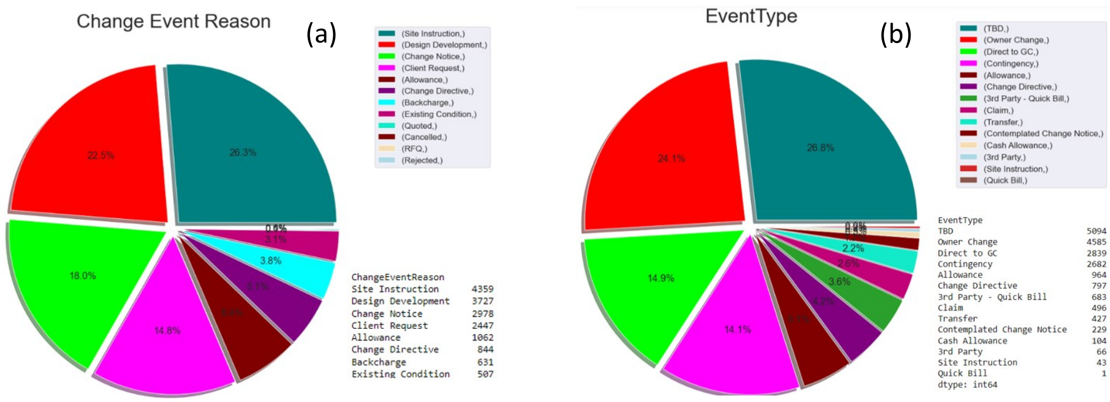


Figure 4-14. Changes events distributions by reason (a) and by type (b)

The contracts and change orders tables contained data from 717 projects. This dataset included, for example, contract value amounts together with the corresponding value amounts of approved change orders per contract for each project. Two main types of change orders were identified in the SQL database according to the contract category of each project: Commitment Change Orders, which represent downstream contractual changes between the general contractor and subcontractors or vendors, and Prime Contract Change Orders, which correspond to formal upstream modifications to the contractual agreement between the project owner (or authorized representatives) and the general contractor. Notably, more than 70% of the recorded change orders belonged to the latter category. It is worth noting that some projects involved more than one contract, and some contracts, in turn, were affected by one or more change orders.

Based on the available data attributes related to both project contracts and approved change orders, an overall change order amount ratio for each project was computed as per (Eq. 6).

$$Change\ Order\ Amount\ Ratio\ \% = \frac{\sum(ApprovedChangeOrderAmounts)}{\sum(ContractAmounts)} \qquad \text{(Eq. 6)}$$

This ratio represents the cost overruns attributable to change orders in each project relative to the original contract value. As shown in Figure 4-15, the average ratio indicates an approximately 19% cost overrun due to approved change orders, and for more than 220 projects this ratio exceeded 20%. These empirical results, derived from real-world project data, underscore the importance of effective change order management in construction projects.

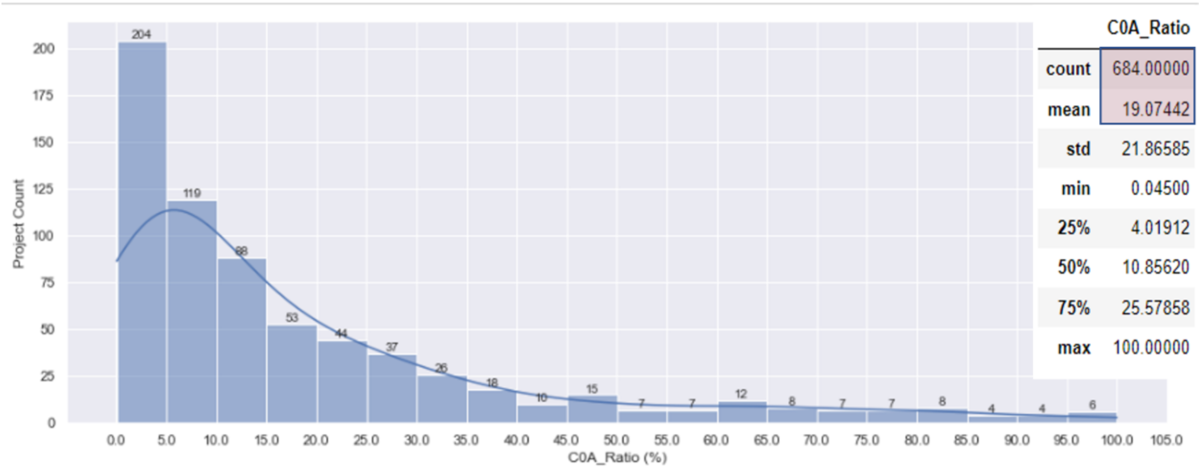


Figure 4-15. Change Order Amount Ratio by Project

4.2.3. Requirement Analysis for Process-Oriented Analysis

To support the requirements analysis for a robust process-oriented approach addressing R02, this study goes beyond identifying and analyzing the data attributes available in the SQL database snapshot of the Change Order Management process (i.e., EDA in previous section). It also examines the overall process structure and its possible routing paths. Since most of the change order data discussed in Section 4.2.2 originates from Procore, Figure 4-16 [203] presents a high-level view of the reference change order approval process. This representation outlines the expected process activities, identifies the responsible stakeholders, and clarifies their intended execution sequence.

As shown in Figure 4-16, the change order approval process typically follows a similar high-level sequence, regardless of the construction management platform used (e.g., Procore) [198, 203]. The process begins with the identification of a change, recorded as a Change Event. If pricing is required, the Change Event is routed to a Request for Quotation (RFQ) subprocess to obtain cost estimates. The proposed changes are then evaluated to determine whether they fall within the original contractual scope or constitute out-of-scope work.

For out-of-scope changes (Prime Contract Change Orders), additional review loops may be introduced prior to final approval, depending on the SOPs (if any) of the AEC/FM organization. Some organizations proceed directly to creating the Prime Contract Change Order and initiate the review and approval stage. Others adopt multi-step workflows, such as a two-tier process

(Potential Change Order (PCO) → Change Order (CO)) or a three-tier process (PCO → Change Order Request (COR) → CO). In this study, the large general contractor managing the project operations data implemented a two-tier process, which is commonly preferred across the industry [198, 203].

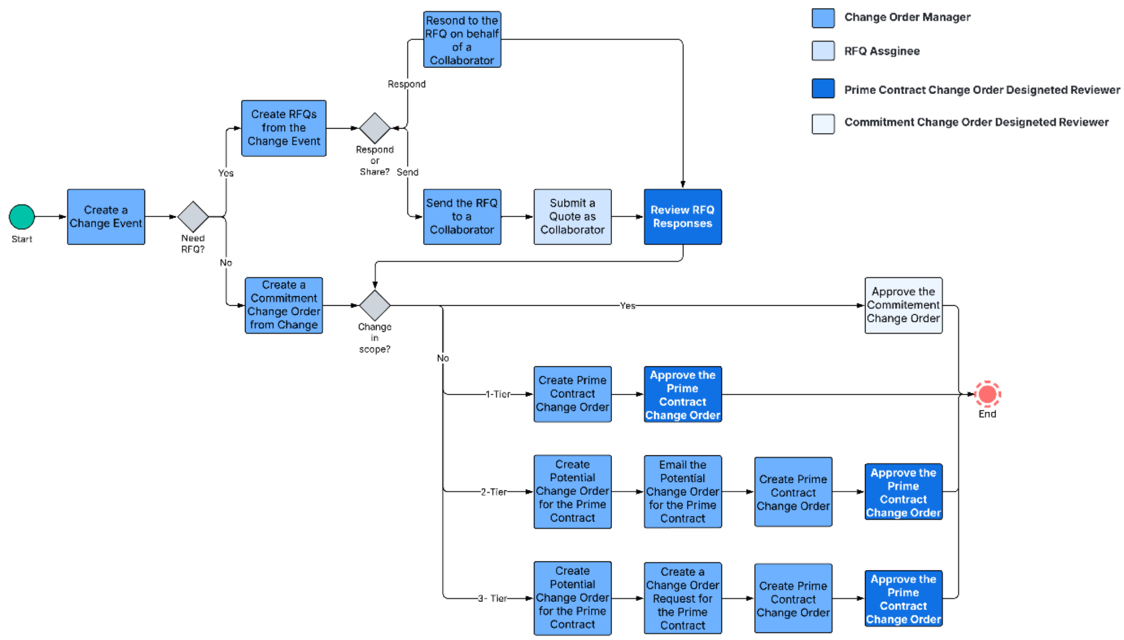


Figure 4-16. High-level reference process model - change order approval (Procore)

Using the high-level process model in Figure 4-16 as a baseline, Table 4-4 was constructed to support both the development of a robust process mining use case for the change order process and the subsequent process-oriented analysis presented in later sections. The table provides an aggregated view of the main process steps (event activities) and maps them to stakeholder roles, data entities, process mining extensions, and corresponding case identifiers (workflow instance IDs).

For example, based on the general contractor (GC) data previously described in 4.2.2, the Create Change Event activity—although it can be initiated by any stakeholder role—is most often created by subcontractors (SCs) within the cloud-based platform. For this activity, it is therefore essential to capture the associated workflow instance ID, execution timestamp, lifecycle status, and related attributes. The same logic applies to all other steps in the change order approval process.

Table 4-4. Process-oriented data overview - Change order approval process

Process Event Name	Actors (Roles)	SQL Source Entity & Extension Name													Case ID				
		Entities						*Extensions*											
		Contracts	Projects	ProjectRole	Workers	ChangeEvents	ChangeOrders	Time	Cost	ID	Lifecycle	Concept	Organizational	Semantic	ChangeEventsID	ChangeOrdersID	ProjectID	ContractID	ProjectRoleId
Identify Project ID	General Contractor (GC) CM , Project Manager (PM) & Owner		✓	✓				✓		✓		✓	✓	✓			✓		✓
Identify Contract ID	GC, CM, PM & Owner	✓	✓	✓				✓		✓		✓	✓				✓	✓	✓
Identify Change Event	GC & Owner			✓	✓	✓				✓		✓	✓		✓		✓		✓
Create Change Event	Subcontractors/ Providers			✓		✓		✓		✓	✓	✓	✓	✓	✓				✓
Create RFQ from Change Event	Subcontractors/ Providers		✓	✓		✓		✓	✓	✓		✓	✓		✓				✓
Create Commitment Change Order	CM&PM			✓			✓	✓	✓	✓	✓	✓	✓	✓		✓			✓
Assign and Send an RFQ to GC	CM & PM			✓		✓		✓	✓	✓		✓				✓			✓
Respond to an RFQ on behalf of GC	PM & Owner			✓	✓	✓		✓	✓	✓		✓			✓				✓
GC Submits Quote/Price	GC/Subs			✓	✓	✓		✓	✓	✓		✓			✓				✓
Review RFQ Responses	CM & GCs/Subs			✓	✓			✓	✓	✓	✓	✓			✓				✓
Approve or Reject Commitment Change Orders	GCs/Subs			✓	✓			✓	✓	✓	✓	✓			✓				✓
Create Prime Contract Change Order	GCs/Subs			✓	✓		✓	✓	✓	✓	✓	✓		✓		✓			✓
Approve or Reject CommitmentChange Order	GCs/Subs			✓	✓		✓	✓	✓	✓		✓				✓			✓
Approve or Reject Prime Contract Change Orders	CM & Owner/PM			✓	✓		✓	✓	✓							✓			✓
Issue Change Order	GCs/Subs			✓	✓		✓	✓	✓	✓	✓	✓		✓		✓			✓
Start Works from corresponding Change	CM & Owner/PM			✓	✓			✓	✓	✓	✓		✓			✓			✓

Although the SQL database snapshot contained rich project operations data related to the change order management process, four critical data completeness challenges affecting process mining implementation were identified: (i) the database snapshot was static, e.g., it did not contain historical lifecycle status records, which are particularly important for analyzing change events and change orders over time and for detecting process drifts; (ii) no linking attribute was found to explicitly connect Change Events with Change Orders in the static dataset; (iii) several data attributes required for a robust process-oriented analysis were missing, including various event activity timestamps, the stakeholder roles responsible for specific process steps, as well as those identified in Table 4-5; and (iv) although the SQL database contained tables related to changes, including projects, contracts, change events, change orders, workers, and project roles, it lacked data for several key subprocesses intertwined with the change order approval process, notably the upstream RFI management subprocess, as well as for downstream subprocesses related to the execution of approved change orders and subsequent progress payments, as illustrated in Figure 4-18 [45, 46, 48].

Table 4-5. Identified missing data attributes in SQL database related to the change order approval process

Source	Source Entity	Attributes
Procore	Change Events	ChangeOrderPackageID; Prime_change_event_id; Commitment_change_event_id; Change_order_request_id. ChangeRequestedBy; RFQId; RFQDateCreated; RFQassginedDate; RFQsubmittedBy; RFQsubmissionDate; RFQRespondedBy; RFQResponseDate; Schedule_impact_amount; SigningDate; SignedBy;
	Change Orders	ChangeOrderRequest ID; ChangeOrderCreatedBy; ChangeOrderReviewedBy; ChangeOrderReviewStartDate. ChangeOrderReviewFinishDate; ChangeOrderApproval; ChangeOrderRejectionDate; ChangeOrderDateIssued; WorflowTierNumber;

4.2.4. Task Mining on Procore's Common Data Environment

Given that the initial static SQL dataset was insufficient to support the development and implementation of process mining for the change order management use case, it became necessary to directly access the primary operational data source, the cloud-based CDE. This environment captures the underlying stakeholders' actions (i.e., HMI) associated with the change order management process and its intertwined subprocesses [46]. In this study, the users' usage data was directly retrieved from the Procore CDE through two complementary mechanisms. First, Procore users' usage reports embedded within the platform were extracted. Second, automated webhooks were configured to capture event-based Create, Read, Update, and Delete (CRUD) transactions, enabled by granted access to multiple large construction general contractors' Procore instances. These webhooks were implemented using Python and the Hypertext Transfer Protocol Representational State Transfer (HTTP REST) API, leveraging multi-factor authentication (MFA), using this endpoint: POST /rest/v2.0/companies/{company_id}/webhooks/hooks.

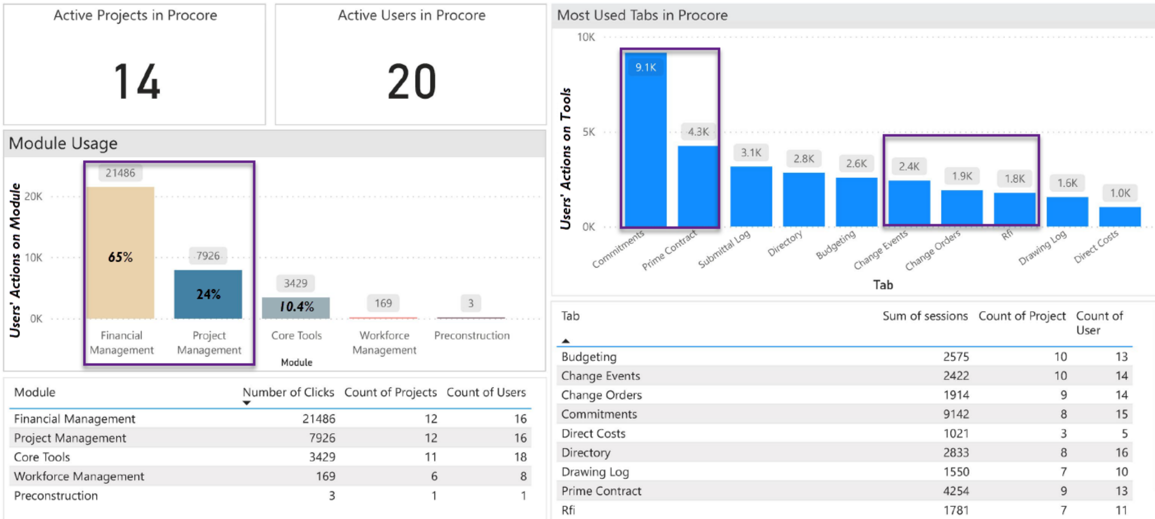


Figure 4-17. Task mining–based analysis of Procore system usage patterns of a general contractor (GC)

As shown in Figure 4-17, this user-centered usage overview served a twofold purpose. Firstly, it enabled the cross-validation of the significance of key business processes by comparing observed user transaction frequencies with findings from the literature review and insights obtained from semi-structured interviews with subject matter experts. This was achieved by analyzing the volume of user transactions performed by project stakeholders across major Procore modules and their associated business process tools. Secondly, for the specific general contractor analyzed in Figure 4-17, the Financial Management and Project Management modules emerged as the most frequently used. Within these modules, tools supporting Commitments, Prime Contracts, Change Orders, Change Events, and Requests for Information (RFIs) were identified as the most actively utilized, confirming their central role in contract administration, change order management, and RFI management processes. These findings are consistent with the processes emphasized by the interviewed industry experts. In general, this user-centered analysis provided a holistic view of the dynamic transactions executed within the CDE, revealing which critical processes were being enacted, by whom, for which projects, and when these actions occurred. Moreover, it helped identify the types of operational project data available within the cloud-based environment that could be systematically extracted via REST API–based data acquisition mechanisms.

Based on the combined insights from industry experts, the reviewed literature, and the explanatory data analysis and task mining results [35, 45-48], a high-level process view was identified to situate RFIs within the broader change order management process, as shown in Figure 4-18. The figure highlights the RFI process as a critical upstream subprocess within a broader network of interconnected workflows involving multiple project stakeholders, including the owner or owner's representative, the general contractor, and subcontractors. RFIs are typically initiated, reviewed, and responded to by the project team, with decision gateways determining whether a change order is required. When such a decision is reached, the process transitions into the formal change order approval subprocess, which primarily involves review and approval activities by the owner or owner's representative. Approved change orders subsequently propagate into intertwined downstream subprocesses, including progress payment processing and the execution of change-related construction work, where subcontractors and trades play a central role and additional verification and approval decision points are introduced [35, 45-48].

Based on this collective evidence, this study focuses its analytical scope on the RFI management process rather than on downstream subprocesses. RFIs represent an early coordination and decision-making mechanism where potential changes may emerge but can also be clarified, mitigated, or resolved before escalating into formal change orders. From a process performance perspective, focusing on RFIs provides an opportunity to identify inefficiencies, communication breakdowns, and delays at an early stage, thereby supporting preventive rather than corrective interventions. While downstream subprocesses, particularly those related to change order approval and execution, remain essential components of the overall system, their behavior is often influenced by upstream information quality and response effectiveness. Consequently, emphasizing the RFI process enables a proactive and analytically justified entry point for process performance assessment, while maintaining the broader change order management context within which these subprocesses operate.

Despite the central role of RFIs and change order management in project delivery, current industry practices lack a systematic, data-driven method for continuously monitoring, assessing, and improving operational process performance across these interconnected workflows. Existing approaches are often fragmented, reactive, and primarily descriptive, relying on static reports or isolated metrics that fail to capture the dynamic and process-oriented nature of construction operations. This limitation constrains organizations' ability to transform increasingly available digital project data into actionable insights that support evidence-based decision making and continuous improvement at the process level.

To address this gap, Chapter 5. introduces and describes the Lean-based Process Mining and Management (LPMM) methodology as a comprehensive framework for data-driven process performance assessment and monitoring. The LPMM framework provides a structured approach to transforming raw operational data into process-aware representations, enabling systematic extraction, validation, and quantitative evaluation of process behavior. By integrating lean principles with process mining techniques, the framework supports objective performance measurement, identification of inefficiencies, and generation of actionable improvement insights, thereby operationalizing digital data for sustained process performance enhancement in construction management contexts.

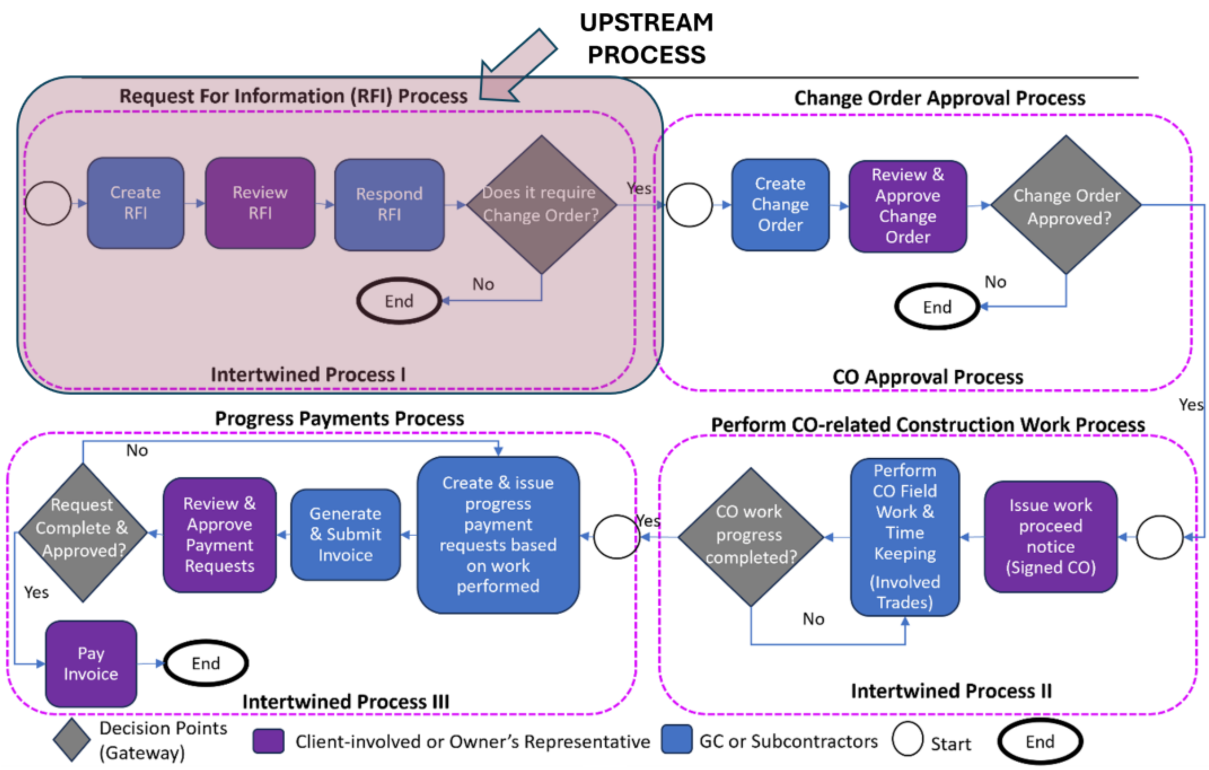


Figure 4-18. High-level process view highlighting the upstream role of RFIs in change order management.

CHAPTER 5. LEAN-BASED PROCESS MINING & MANAGEMENT (LPMM) METHODOLOGICAL FRAMEWORK[4]

5.1. Introduction

Traditional construction management practices, although well established, focus primarily at the project level on planning and controlling schedule, cost, and risk; ensuring technical feasibility and compliance; and managing construction methods and labor resources. Unlike this traditional approach that tends to overlook core business processes, project-driven organizations with strong process awareness have a clear advantage over competitors, as they continuously monitor and control the performance of their business operations. This allows them to execute and manage their recurring processes more efficiently each time, sustain long-term effectiveness, and learn from each project along the way to further enhance their future operational performance [204, 205]. Construction projects, ranging from low to high complexity, often involve multiple company stakeholders requiring well-established intra- and inter-organizational processes to ensure seamless collaboration and an efficient project realization [25, 164]. These processes are the backbone of successful project management, streamlining day-to-day operations, enhancing coordination, and ultimately improving overall project performance. Construction organizations can operationalize their business strategies, optimize workflows, reduce inefficiencies, and make informed decisions by implementing well-structured processes and actively managing their dynamic performance [26, 93, 128, 169].

As highlighted in Gap 1 of Chapter 3. , most means and methods currently used in the construction domain for extracting corporate knowledge on operational business processes, such as workshops, interviews, and focus groups, are manual, time-consuming, and subjective [75]. Although these methods assist construction organizations in conceptually documenting and structuring their baseline business processes, the resulting process models are often static, resource-intensive, and error-prone. Traditional ontology-based methods offer a more formalized and machine-interpretable approach, enabling semantic consistency, reusability, and automation

[4] This chapter is based upon:
A. J. Martinez Lagunas and M. Nik-Bakht. 2025. " An Automated Data-to-Process (D2P) transformation method for RFI process performance assessment and monitoring". Status: Submitted – Under Review

potential. However, they still heavily rely on tacit expert input and may become outdated if not continuously maintained or enriched [24, 160]. What's more, these methods alone fail to uncover and diagnose evidence-based process performance issues dynamically and automatically.

In contrast, data-driven methods such as process mining can leverage process-related data generated by various stakeholders across multiple management Information Systems (IS) to automatically discover, monitor, and optimize 2E2 business processes. This approach provides a dynamic and fact-based view of the actual processes' executions and how they actually perform, enabling continuous process improvement and AI-driven automation opportunities [31]. This is particularly important in the construction industry not only because projects are process-intensive, but also because these processes are often executed differently than expected or planned, change over time, and are subjected to various uncertainties. Considering these aspects and recognizing the globally documented applications of process mining across various industries and the widely reported benefits of this technology for enhancing business operations [8, 9], scholars and industry practitioners are increasingly exploring its potential in the construction domain [26, 163]. Nevertheless, existing data architectures of most construction organizations lack process-awareness, which hinders both the implementation of process mining and its integration with other process-oriented management methodologies for analyzing the performance of core construction business operations [26]. AEC/FM organizations often fail to harness data from their day-to-day operations to extract fact-based, process-oriented performance insights [46, 206, 207]. One of the main factors hindering the broader adoption and implementation of this technology is the lack of a data-driven, process-oriented method tailored to the construction domain, one that can shed light on how to collect and transform raw operations data from construction PMIS to standard data schemas suitable for process mining implementation and dynamic process performance monitoring. In this regard, this study seeks to answer two main research questions: (RQ1) How can construction organizations harness project operations data and extend their existing data-oriented architectures to become process-aware? (RQ2) How can process mining techniques be implemented and integrated with Lean-based principles and metrics [208] to automatically extract, quantitatively assess, and continuously monitor the actual E2E performance of construction business operations with evidence-based actionable insights?

To tackle these challenges and address these questions, this study proposes a novel data-driven, Lean-based Process Mining and Management (LPMM) methodological framework tailored to the construction domain that integrates process mining techniques with Lean-based metrics. This hybrid approach will facilitate dynamic performance monitoring of core construction business processes. Furthermore, this work aims to drive process intelligence (PI) in construction projects by leveraging advanced technologies to automatically extract process-oriented knowledge, including actionable process performance insights, to support data-driven decision-making by key stakeholders such as owners, project managers, project coordinators, and project controllers, thereby enhancing overall construction process management. To achieve these goals, the underlying objectives are to: (i) develop a data-driven LPMM framework tailored to the construction domain to facilitate automated process modeling and dynamic Process Health Monitoring (PHM), regardless of the specific IS or Common Data Environments (CDEs) used to manage a construction project (i.e., system-agnostic), and applicable to any critical business process (i.e., process-agnostic); (ii) design a process mining use case for analyzing significant inter-organizational processes in real-world construction projects; and (iii) develop the required algorithms to enable the automated generation of event data logs necessary for building Foundational Process Data Models (FPDMs) to support process mining implementation. This includes the connector, extractor, serializer, and logger components, while ensuring that the generated event logs conform to the eXtensible Event Stream (XES) standard [108]; (iv) apply the proposed LPMM methodological framework in real-world construction projects to quantitatively assess and monitor process performance; and (v) analyze and compare the performance of discovered process executions across similar construction projects, while identifying data-driven strategies for process improvement and automation based on the extracted performance insights.

The proposed data-driven process mining LPMM framework, presented in Figure 5-1 developed in this study is grounded in [50, 209, 210]. This systematic methodology aims to assist construction organizations in harnessing project data generated from day-to-day operations to automatically reconstruct actual process executions, diagnose major operational inefficiencies, and quantitatively assess process performance. Moreover, this novel process-oriented framework provides construction practitioners and scholars with a data-driven methodology to guide and streamline the implementation of process mining for the analysis and continuous monitoring of "as-happened" construction business processes.

5.2. Data-To-Process (D2P) Transformation Method

The two-layer LPMM framework, introduced in Figure 5-1, is structured as a five-level stepwise progression, in which each level incrementally builds upon the preceding one. The data-oriented architecture (Layer I) comprises two foundational levels, namely systems' raw data and project-centric contextual data, reflecting typical information architectures in medium- to large-scale AEC/FM organizations that lack process awareness. To enable this capability, the process-oriented structure (Layer II) consists of three upper transformational levels that support the development of foundational process data models, enable automated process extraction and performance assessment, and facilitate evidence-based decision-making to enhance process control and redesign for sustained improvement through the Data-to-Process (D2P) transformation method. Furthermore, the decision risk level will normally depend on the degree of process-awareness enabled, the higher the data-driven process-awareness the lower the decision risk as depicted in Figure 5-1.

Layer II of the LPMM framework comprises the D2P method, which includes seven core process-aware modules that collectively enable the systematic extension and transformation of conventional data-oriented architectures into fully process-aware structures. Module 1 of the D2P method establishes the data-driven foundation and requirement analysis necessary for process mining implementation by automating the extraction, integration, and preprocessing of heterogeneous project-centric data from CDEs and IS, supported by exploratory data analysis, task mining, and expert consultation. Module 2 comprises the process mining use case development and the definition of its analytical scope through the formulation of key analytical questions, process model perspectives or MVDs, key PPIs, abstraction levels, and stakeholder viewpoints, informed by a systematic literature review and aggregated domain expert input. Module 3 enables automated event log construction by transforming integrated datasets into standardized, process-aware representations through the algorithmic development of automated data extractors, process-oriented schema transformers compliant with the eXtensible Event Stream (XES) standard, logging components, and data enrichment pipelines. Building upon this foundation, Module 4 applies process mining techniques, including Fuzzy Miner and Inductive Miner, to automatically discover E2E "as-happened" process model executions. Module 5 derives process efficiency formulations grounded in process mining theory and applies quantitative performance metrics across real-world construction projects to evaluate efficiency, detect deviations, analyze process variants, and assess the influence of RFI content on variant behavior using NLP. Finally, Modules 6 and 7 translate these analytical insights into actionable process improvement strategies and enable continuous PHM through purpose-built process performance boards supporting key business operations in construction projects.

Grounded in the principles of scalability, reusability, and automation, the proposed D2P method enables systematic cross-project deployment, supports architectural extensibility across inter-organizational business processes, and streamlines targeted evidence-based interventions for process redesign and continuous process improvement.

5.3. Layer I of the LPMM Framework: Data-Oriented Architecture

To stay competitive and fulfill contractual obligations with tight deadlines, most construction companies support their business operations by implementing various technologies such as PMIS, ERP systems, CDEs, construction management systems, accounting software, or digital databases. AEC/FM organizations often implement one or more of these data-oriented digital systems as part of their operational data strategy. These platforms serve as support tools to manage construction projects across their lifecycle phases in a more structured and efficient manner.

As shown in Figure 5-1, the first stage (Layer I) of the proposed D2P transformation method involves the systematic identification of the foundational components of the existing data-oriented architecture layer. including the primary data agents (i.e., source systems) deployed by AEC/FM organizations and the key project data attributes associated with critical business processes. This data-oriented architecture (Layer I) typically comprises two primary levels: (L1) the system-level raw data, consisting of heterogeneous, non-contextual operational records., e.g., Stakeholder-To-System interactions or transactions performed across multiple construction digital platforms; and (L2) the project information level, comprising structured and contextually enriched datasets prepared to enable project-level analytical performance monitoring and managerial decision-making.

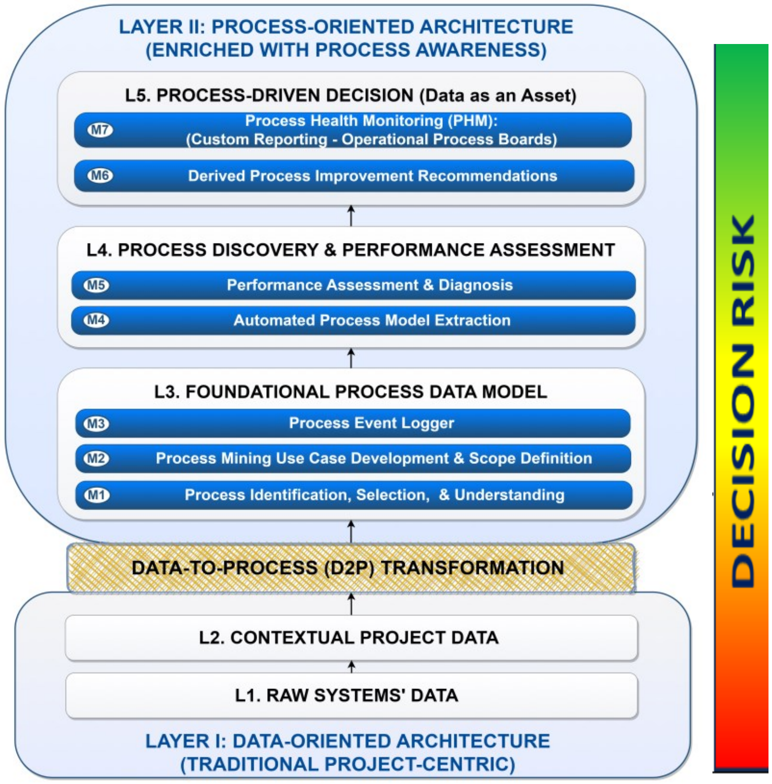


Figure 5-1. Lean-based Process Mining and Management (LPMM) framework: D2P for automated PHM

5.3.1. Data-oriented systems, project roles, and raw data – Level 1 (L1)

At the foundational level (L1), the D2P method systematically maps the data-oriented system architecture implemented by AEC/FM organizations, identifies the primary business processes supported by each construction digital platform, analyzes stakeholder–role interactions across operational functions, examines the underlying raw data structures, and evaluates existing cross-system data exchange mechanisms, where applicable. The identified raw project operations data is leveraged at later stages to produce valuable project information, extract actionable data-driven insights, and build corporate knowledge assets. Some widely adopted commercial cloud-based system platforms for managing construction projects in North America include, among others: Procore [199], Aconex [211], Autodesk Construction Cloud (ACC) [212], Oracle NetSuite [200], System Applications and Products (SAP) [213], CMIC [214], Sage [215], and QuickBooks [216]. Typically, construction companies adopt one or more of these digital platforms, alternative solutions, or internally developed systems, depending on their operational requirements, project-specific client expectations, and overarching document and data management strategies.

Regardless of the selected system(s), these digital platforms (i.e., project-centric CDEs) often contain large volumes of valuable raw data and project information that often remain underutilized and not fully leveraged. In this context, task mining, a data-driven technique, is applied to systematically examine Stakeholder-To-System interactions with the actions performed on the IS or CDEs to extract the task-level execution patterns related to significant inter-organizational construction business processes [35, 71]. This study assumes the existence of digital project operations data stored within construction digital platforms deployed across organizational project portfolios to support construction business operations.

5.3.2. Data processing, data model, and project information – Level 2 (L2)

The second level of the proposed D2P framework consists of understanding the backend system model architecture and procedures implemented by AEC/FM organizations to store and manage their projects' data related to their business operations. Unlike L1, Level 2 (L2) additionally involves identifying the data context, including how project operations data is stored and how data attributes are related across different construction project entities. Different IS provide specific features and capabilities that support specific business operations across the lifecycle of a construction project. This aspect often forces companies to implement more than one of these systems to meet project requirements, leading to a major challenge as their operations data and workflows are hosted across multiple source systems in silos. To tackle this challenge, medium to large organizations implement different technological solutions such as data lakes with data factories [217] as part of their data-architectures to extract, integrate, store and process their projects' data from various source systems. This architecture enables organizations to structure their operations data, often by constructing Common Data Models (CDMs), that is, a shared language and structure for data with machine-readable schemas to support business operations. CDMs provide context, syntax, and semantics to business entities and their corresponding attributes, thereby facilitating the understanding and standardization of operations data originating from multiple source systems. CDMs serve as a foundation to create system-specific Entity Relationship Diagrams (ERD) that often include multiple data entities, attributes and table relationships for projects being managed in that specific system. The ERD defines the logical data schema and its relational organization, typically implemented within Structured Query Language

(SQL) relational database management systems hosted on cloud-based infrastructures or on-premises SQL servers [218].

Web-based construction project management platforms, CDEs and IS, often incorporate document management modules that store project artifacts such as contracts, schedules, drawings, invoices, RFIs, BIM models, field reports, change orders, and issue daily logs. These artifacts are typically indexed within relational database management systems (e.g., SQL database) as structured data attributes linked to associated digital artifacts. When automated and recurrent data pipelines are properly configured, these digital construction platforms enable systematic extraction of both historical and up-to-date contextual project operations data. Collectively, such data-oriented IT architectures constitute the analytical backbone required to generate advanced data-driven performance analytics at the project-level, thereby enabling evidence-based decision-making and robust multi-dimensional project control across cost, schedule, quality, and other critical performance dimensions. However, from a methodological standpoint, the first two staging levels of data-oriented IT architectures depicted in Figure 5-1 remain insufficient, as they are predominantly project-centric and lack explicit process awareness necessary for process mining–based analysis. In practice, many AEC/FM organizations operate primarily at this project information-centric level (L2), without systematically capturing or analyzing the execution performance of inter- and intra-organizational business processes. Consequently, these data structures do not natively represent process flows, control-flow dependencies, or cross-functional interactions. This limitation is critical, as poorly managed or inadequately monitored business processes can significantly compromise overall project performance across cost, schedule, operational efficiency, and other key performance dimensions.

5.4. Layer II - LPMM Framework: D2P Method - Process-aware architecture

To address the aforementioned structural limitation, the proposed D2P method introduces a process-oriented architectural layer that reconstructs actual business process execution sequences from digital project operations data, thereby enabling automated modeling, monitoring, and diagnostic performance evaluation of E2E ‘as-happened’ process executions. Through this transformation, existing data-oriented architectures are extended into process-aware analytical constructs capable of supporting continuous, multi-perspective process performance monitoring.

These constructs enable the systematic extraction of actionable process knowledge assets that inform evidence-based decision-making and facilitate timely, targeted interventions to enhance operational efficiency and control process variability. To embed process-aware capabilities within existing data-oriented architectures, the process-oriented layer of the D2P method is structured into three incremental levels comprising seven interdependent modules, as depicted in Figure 5-1. Further details regarding their methodological specification are presented below.

5.4.1. Foundational Process Data Models (FPDMs) – Level 3 (L3)

Level 3 focuses on the construction of foundational process-oriented data models in the form of structured event logs derived from project operational data distributed across heterogeneous and often siloed information systems (IS) and digital relational databases. At this stage, a target inter-organizational process is selected based on defined analytical criteria, including operational relevance, data completeness, and its role within the broader process ecosystem. Process event data logs are then systematically constructed to capture E2E 'as-happened' process behavior. These event logs serve as the formal analytical foundation for the application of process mining techniques and the generation of actionable, performance-centric process knowledge assets. Model development entails: (i) structured requirement analysis and process understanding; (ii) formalization of the process mining use case and analytical scope; and (iii) algorithmic development for the automated generation and validation of high-quality event logs, together with formally defined key process performance metrics, enabling quantitative analysis of process variations and inter-organizational execution patterns to extract evidence-based process performance insights.

5.4.1.1. Module 1 (M1): Process and project selection, process understanding

Module 1 at Level 3 (L3) of the D2P method comprises project and process selection, along with the understanding of the selected process for process mining implementation. While process instance volume (i.e., the number of executed instances of the selected process within the CDE or IS) constitutes a primary requirement for process mining–based analysis in real construction projects, project selection must also account for additional criteria. These include data availability, data completeness, and project-specific characteristics, such as the delivery system, project type, and contract type.

Module 1 (M1) establishes the foundational process mining requirements for the process-based cross-project analysis through integrated knowledge capture and project operational data acquisition. Domain expertise is systematically captured and analyzed to validate the operational relevance of the target inter-organizational processes and to formalize a reference process model commonly implemented in AEC/FM projects (see Figure 4-1). Such significant business processes are often executed across multiple stages of the project lifecycle, reflecting their cross-phase operational relevance and their potential to influence downstream project control processes.

In terms of process understanding, the implementation stage of the D2P method (6.2.1.1) formally specifies the selected process as typically executed in real-world construction projects, drawing on project operational data and established industry practices to define a reference 'as-planned' process model. This includes the identification of core process activities, decision points, involved stakeholders, and inter-organizational interactions, together with the required information exchanges. This process-oriented module (Module 1) also integrates outputs from the data-oriented architectural layer (see Figure 5-1), including task-based analysis, to identify and formalize the complete set of project operational data required for process mining–based investigation. This structured data is subsequently extracted and transformed in later D2P stages to enable E2E process performance assessment.

Despite their operational importance, inter-organizational processes in construction projects remain difficult to monitor dynamically across project portfolios. Prior studies [26, 223] suggest that the integration of process mining with Lean principles enables automated, objective, and continuous performance assessment. However, this capability remains largely unrealized in practice due to the lack of process awareness in existing construction data architectures. Consequently, organizations are unable to derive actionable, data-driven insights into actual E2E process performance. This limitation is addressed in this study through the implementation of the proposed LPMM methodology.

5.4.1.2. Module 2 (M2): Process mining use case development and scope definition

Module 2 (M2) formalizes the process mining use case and defines the analytical scope for process performance assessment. Building upon the outcomes of M1, it specifies the primary Common Data Environments (CDEs) and information systems (IS) used by inter-organizational stakeholders to manage project operations and execute the target process. It further consolidates the identified Human–Machine Interactions (HMIs) performed by each stakeholder role across digital platforms to operationalize the process. The resulting extracted and integrated process-related data form the required input for event log generation. These event logs serve as the operational backbone for reconstructing 'as-happened' process executions and deriving data-driven performance insights through process mining techniques.

The process mining use case developed in this study, outlined in Section 6.2.1.2, comprises three core tasks: (i) automated process extraction, (ii) data-driven performance assessment, and (iii) continuous Process Health Monitoring (PHM) through customized process monitoring reports. The analytical scope spans four main process mining perspectives, defined as process model view definitions (MVDs): control-flow, cost, time, and organizational. The *control-flow* perspective examines process execution variants, including their sequence patterns and frequencies; the *time* perspective reveals the execution times for different parts of the process and for the overall E2E process; the *cost* perspective evaluates E2E activity-based costs across process executions; and the *organizational* perspective analyzes inter-organizational role behavior and its impact on E2E process performance.

Table 5-1. PHM Scope Definition & Key Process Performance Metrics: Control-flow & Time Perspectives

Perspective	Type of Analytics	Process-Question(s) Base Investigation Scope	Key Process Performance Metric(s) & Refs.
Control-flow (CF)	Process Discovery & Variant Analysis (Descriptive)	CF-Q1. What is the executed dominant process model? CF-Q2. What are the most common process variants?	Fitness; Precision; Frequency; Rework Loops [90, 121-123, 219]
	Conformance Checking & Variant Analysis (Diagnostics)	CF-Q3. What are the process deviations & frequency-based bottlenecks? CF-Q4. Which process variants exhibit the highest level of rework? CF-Q5. Why did the process deviations, bottlenecks, or rework happen?	
Time (T)	Performance Assessment (Descriptive)	T- Q1. What is the inter-arrival rate? T- Q2. What percentage of on-time process executions? T- Q3. What is the Weighted Average Process Efficiency (WAPE)? T- Q4. What process variants had the highest average lead times?	Inter-arrival rate; % of on-time and late responses WAPE [220-222]

	Performance Assessment & Root Cause Analysis (Diagnostics)	T- Q5. Which process event activities are most associated with time-related bottlenecks? T- Q6. Are there any observed process drifts or batching behaviors? T- Q7. To what extent do rework loops contribute to increased process lead times? T- Q8. What process execution types have longer activity cycle times? T- Q9. When and where in the process do work peaks occur? T- Q10. What observed factors or patterns explain the variation in process time efficiency between the two investigated projects?	Cycle Time; Lead Time; Work In Progress (WIP); Performance Spectrum; Process Efficiency [223-226]

Module 2 (M2) includes scope definition through the specification of process-oriented questions to be addressed in the data-driven process performance assessment, as presented in Table 5-1 and Table **5-2**, which outline the analytical techniques and their associated key process performance metrics across each perspective. Building upon established metrics (see references in Table 5-1 and Table **5-2**), this study formalizes time- and cost-based process efficiency measures grounded in data-driven process mining evidence, directly addressing Gaps (4) and (6) identified in Section 3.6.

Table 5-2.PHM Scope Definition & Key Process Performance Metrics: Cost & Organizational Perspectives

Perspective	Type of Analytics	Process-Question(s) Base Investigation Scope	Key Process Performance Metric(s) & Refs.
Cost *(C)*	Performance Assessment (Descriptive)	C-Q1. Number of process instances per million dollar of construction value? C-Q2. What is the average cost of the discovered dominant process? C-Q3. What is the total cumulative process cost incurred? C-Q4. Which are the costliest process variants? C-Q5. What is the Weighted Average Cost Efficiency (WACE)?	Process Instances/$Million Cost Amount; Cost Overruns Amount; Cost Efficiency (WACE) [227, 228]
	Performance Assessment & Root Cause Analysis (Diagnostics)	C-Q6. What is the average cost overrun attributable to rework loops? C-Q7. What are the primary process execution drivers contributing to cost overruns? C-Q8. What are the most cost-efficient process variants? C-Q9. What process characteristics lead to higher processing costs?	
Organizational *(O)*	Social Network Analysis (SNA) (Descriptive)	O-Q1. Which resources are frequently involved in sequential task handovers? O-Q2. Which resources co-participate most frequently in the process?	SNA Metrics for: Handover of Work; Working Together; Subcontracting

	O-Q3. What are the subcontracting patterns among process participants?	[90, 229-231]
SNA Performance Assessment (Diagnostics)	O-Q4. Do overloaded resources experience longer handling times? O-Q5. Do specific process participants or roles influence process delays?	

5.4.1.3. Module 3 (M3): Process event log generator

AEC/FM organizations constantly strive to harness project operations data from siloed source systems in order to enable the automated extraction of actual process performance insights. This task requires the integration and transformation of raw system data into well-defined, process-aware structures [232, 233]. Currently, the AEC/FM domain lacks a systematic method to accomplish this task. To address this gap, the D2P transformation method comprises a system-agnostic module for automated process event log generation, illustrated by Figure 5-2, which follows a data Extraction, Transformation, Loading, and Enrichment (ETL+E) pipeline to construct high-quality process event logs.

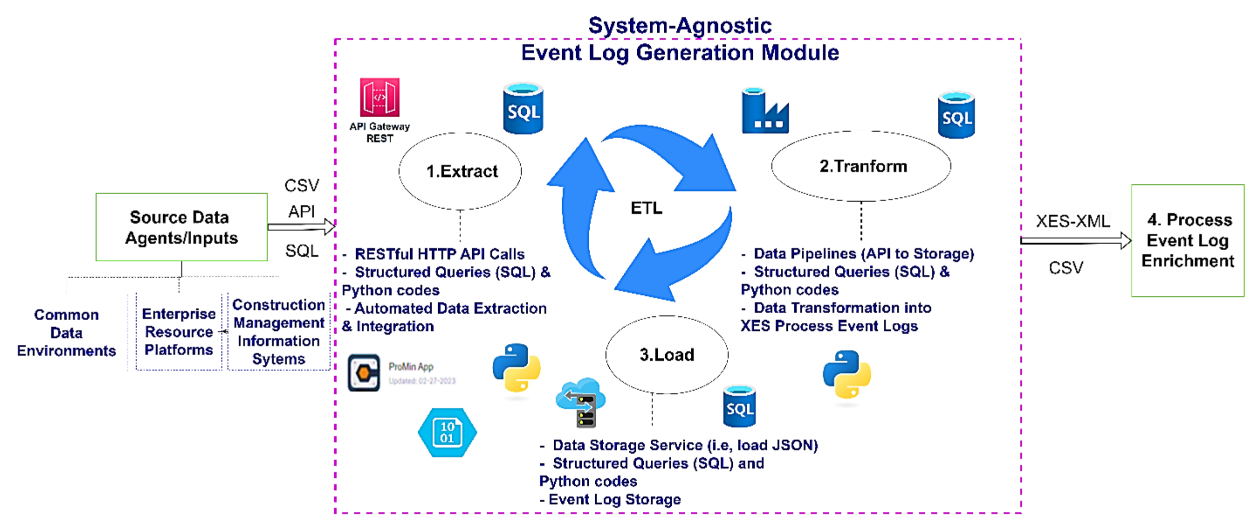


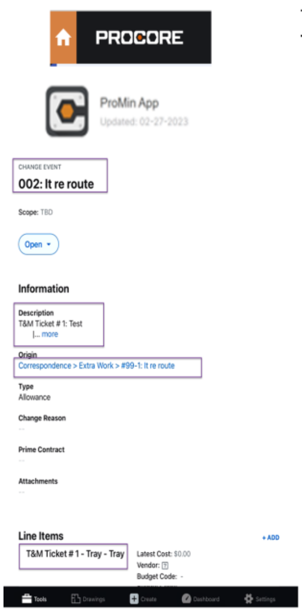
Figure 5-2. System-agnostic event log generation module.

The data extraction phase involves harvesting project operational data from multiple heterogeneous sources, including relational databases (e.g., SQL-based systems) and construction digital platforms such as CDEs. Extraction from relational databases is typically performed using structured SQL queries, with outputs exported in Comma-Separated Values (CSV) format. However, project data stored in such databases is often available only as static snapshots captured at discrete points in time, thereby lacking explicit E2E process lifecycle awareness.

To address this limitation, in this Module 3 (M3) additional data extraction mechanisms are implemented to retrieve E2E process-related attributes directly from cloud-based construction platforms (i.e., CDEs). This extraction is performed through HTTP-based REST (Representational State Transfer) APIs. In this vein, Algorithm 1 presents the developed data extractor, illustrated in Figure 5-3 is defined to systematically retrieve for example RFI process-related data, including lifecycle status records, role-based project actions, and associated timestamps, from Procore. Although the D2P method is demonstrated in this study using Procore as the primary project CDE [199, 234], the underlying data extraction logic is generalizable to other CDE platforms and process types by targeting their respective API endpoints.

In this study, the proposed extraction algorithm is configured to support recurrent, automated retrieval of process data from the CDE. First, a secure Data Connection Web Application is registered through the Developer Portal [235] to enable API-based integration. Second, the extraction logic is implemented as an HTTP-triggered Azure Function App [217], authenticated via bearer token credentials, and configured to target two principal process API endpoints: (i) */projects/PROJECT_ID/rfis/* for retrieving structured process metadata and content, and (ii) */projects/PROJECT_ID/rfis/RFI_ID/replies* for capturing associated process response threads and inter-organizational information exchanges and communications. Third, a cloud-based object storage resource is provisioned to persist the raw JSON (JavaScript Object Notation) response payloads, thereby ensuring data traceability, reproducibility, and auditability of the extraction process. Fourth, the nested JSON structures are programmatically flattened and schema-aligned to extract relevant RFI attributes, including unique identifiers, temporal markers, workflow states, responsible actors, and inter-organizational communication exchanges. Fifth, SQL stored procedures are configured to transform and load the extracted JSON data into normalized relational database schemas, enabling structured querying and longitudinal analysis. Sixth, the Data Extractor is deployed within a cloud-based execution environment and configured to operate on a recurrent schedule to ensure continuous and systematic acquisition of process data. Finally, a data fusion mechanism reconciles newly retrieved RFI process-related records with pre-existing relational datasets to maintain dataset integrity and preserve historical state transitions. This integration step ensures data consistency while producing a consolidated, lifecycle-aware dataset that serves as the structured and temporally ordered foundation for automated process event log generation via the D2P transformation method.

Inputs	Developed Algorithm	Outputs



Development of a Procore Data Connection App

❖ API Extractor Algorithm

Algorithm 1: RFIs DATA EXTRACTOR FROM REST API

```
input  : COMPANY_ID                         // Company identifier
         PROJECT_ID                         // Target project ID
         token                  // Bearer token for authentication
output: rfiList, RepliesList    // RFIs and their replies for the given
         project
begin
    BASE_URL ← https://api.procore.com;
    endpoint_1 ← /rest/v1.0/projects?company_id= COMPANY_ID;
    headers ← Authorization:  Bearer  token;
    url_1 ← BASE_URL + endpoint_1;
    response_1 ← SendHTTPRequestGET, url_1, headers;
    projectList ← ParseJSONresponse_1;
    foreach p ∈ projectList do
        if p.id = PROJECT_ID then
            endpoint_2 ← /rest/v1.0/projects/ +PROJECT_ID+
             /rfis;
            url_2 ← BASE_URL + endpoint_2;
            response_2 ← SendHTTPRequestGET, url_2, headers;
            rfiList ← ParseJSONresponse_2;
            foreach r ∈ rfiList do
                RFI_ID ← r.id;
                endpoint_3 ← /rest/v1.0/projects/ +PROJECT_ID+
                 /rfis/ +RFI_ID+ /replies;
                url_3 ← BASE_URL + endpoint_3;
                response_3 ← SendHTTPRequestGET, url_3, headers;
                replies ← ParseJSONresponse_3;
                AppendRepliesList, replies;
    return (rfiList, RepliesList);
```

Data Fusion in SQL Database (i.e., with extracted stateful data and timestamps)

Figure 5-3. Pseudocode: Web-based RFI data extractor via HTTP REST API.

Process event logs comprise structured collections of events that capture valuable digital traces, or footprints, of historical executions of construction processes. An event is composed of at least three base attributes: (i) ‘cases’, referring to individual process instances represented by a collection of related events; (ii) ‘event activities’, referring to discrete process actions or transactions recorded at specific points in time; and (iii) ‘timestamps’, referring to the temporal attributes of executed event activities formatted in accordance with the ISO 8601 standard [91, 126]. To ensure robust process mining implementation, event logs must be systematically structured with formally defined syntax and semantics in compliance with established international standards, such as the eXtensible Event Stream (XES) specification, encoded using an Extensible Markup Language (XML) Schema Definition (XSD) [236].

In traditional process mining, events are assumed to be partially ordered, meaning that they follow a strict partial order that is transitive (If a≺b and b≺c, then a≺c), irreflexive (a ⊀a), and asymmetric (If a≺b, then b⊀a). An event log can be formally defined by (Eq. 1) [91]. For a common control-flow perspective analysis, this event log notion can be further simplified as per (Eq. 2) [91] given that only the activities' key-value pairs need to be sequentially ordered for a given process workflow instance. To build this structure, Figure 5-4 presents the developed Transformer (Algorithm 2) as part of the D2P method, which for instance operationalizes the generation of RFI event logs by systematically transforming project-centric, data-oriented schemas (e.g., SQL databases) into structured, process-aware event log representations, illustrating the transformation logic for the 'Create RFI' process event activity and its applicability to other process event activities within the E2E process subject to assessment and monitoring.

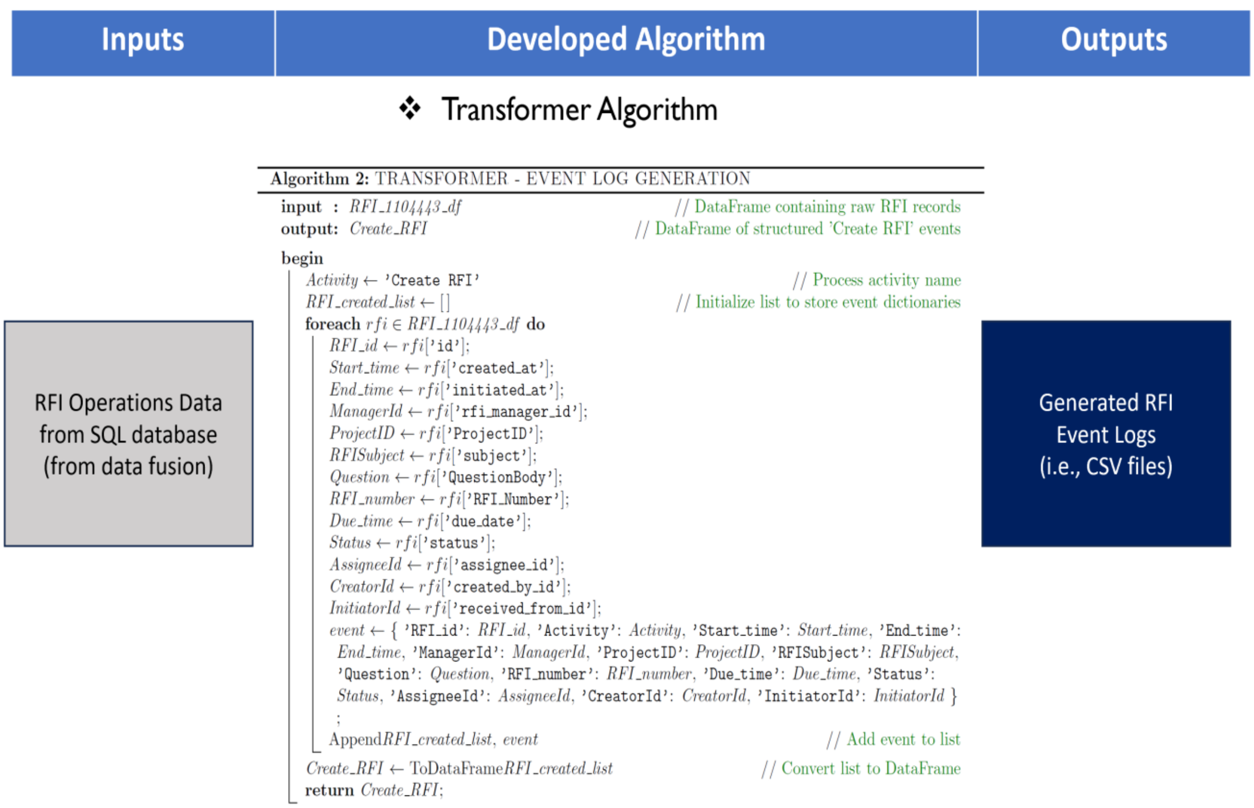


Figure 5-4. Pseudocode: Transformer for automated event log generation.

Once constructed, the event log structure can be enriched to enable additional perspectives of analysis, and subsequently serialized into XES format to adhere to the standard, for instance, by using the python library PM4Py [98, 207]. Figure 5-5 presents an example of event log enrichment by adding the cost extension (or dimension) to the RFI process event log, thereby enabling analysis of the process from a cost perspective. Additionally, the algorithm integrates the XES serializer and a loader mechanism for storing the enriched event log back into the SQL database [207], as illustrated in Figure 5-5.

The enrichment algorithm (Algorithm 3) comprises three main steps. First, a web-based connection to the SQL server is established, and the event log data table is retrieved using structured SQL queries. Second, for each process event activity in the log, activity-based cost values are computed per project stakeholder role and appended as cost-related attributes. Third, the enriched RFI event log is serialized in conformance with the XES schema definition to ensure structural and semantic compliance and subsequently persisted back into the SQL server database.

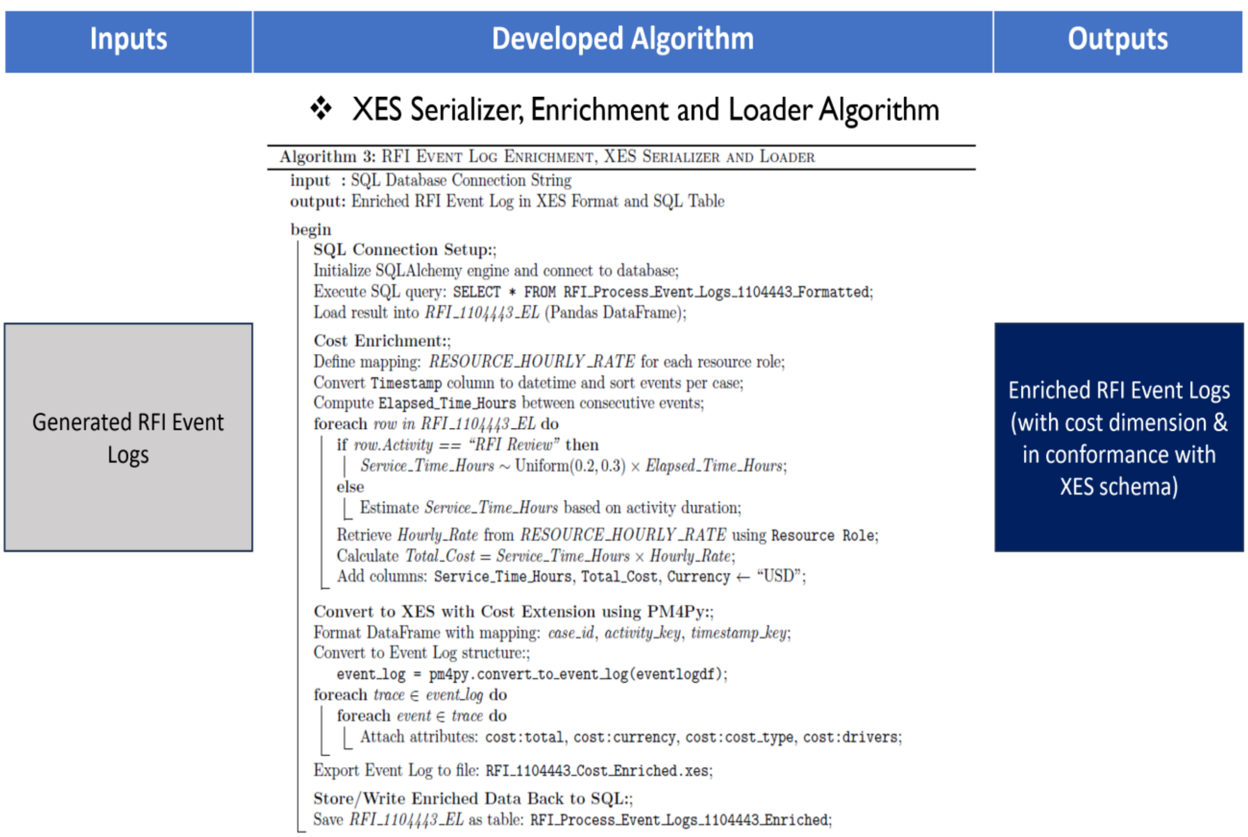


Figure 5-5. Pseudocode for RFI event log enrichment and XES serializer.

5.4.2. Process discovery and performance assessment – Level 4 (L4)

5.4.2.1. Module 4 (M4): Process model discovery/extraction

Event logs, produced by algorithms similar to those previously presented in Section 5.4.1.3, serve as the basis for implementing process mining techniques [35]. This module focuses on the automated process discovery and reconstruction of actual process model executions. Process mining enables this capability, whether to extract "as-happened," near-real-time, or real-time process models, depending on the implemented data extraction and transformation techniques as well as the level of process awareness, as illustrated in Figure 5-1.

The α-algorithm often proves ineffective to discover real-world process behavior due to its inability to filter out non-significant behavior, but it serves as the foundation for more advanced process discovery techniques, such as the Fuzzy Miner and Inductive Miner (IM) algorithms [121-123]. These process discovery algorithms, as explained in Section 2.2.5, are widely used in real-world industrial applications, as they support the extraction of more realistic and interpretable process models by filtering out non-significant execution paths (i.e., Fuzzy Miner) and ensuring soundness or the absence of deadlocks (i.e., IM), which allows the E2E execution of each process instance. Therefore, this study adopts these two techniques to automatically derive the “as-happened” inter-organizational management process model from constructed operational event logs and enable continuous process performance monitoring.

To perform a robust process performance analysis, not only must the event logs conform to existing standards (i.e., XES), but the quality of the discovered process models must also be evaluated. To perform this verification this study relies on two main quality metrics proposed by Buijs et al., (2014): (i) *‘fitness’*; and (ii) *‘precision’*. Both of these metrics work by replaying the observed behavior of the event log on top of the process model discovered using an alignment-based technique [126]. However, the former analyzes the fraction of the traces in the event log that can be reproduced by the process model by and it can be computed using (Eq. 7) [237]. In other words, it measures the extent on how well the extracted process model properly reproduces the observed behavior in the event log. The latter, instead, focuses in analyzing any additional behavior allowed by the process model that is not present in the event log by counting the so-called escaping edges on the model for each aligned trace event (i.e., visits), as per (Eq. 8) [237].

$$Fitness = 1 - \left(\frac{\sum_{logtrace-i}^{logtrace-\boldsymbol{n}}(f_{AlignCostMismatches}(L, M))}{\sum_{logtrace-i}^{logtrace-\boldsymbol{n}}(C_{L_{Non-SynchonousMoves}})}\right) \quad \text{(Eq. 7)}$$

Where $f_{AlignCostMismatches}(L, M)$ represents a cost function with the 'penalty' given to the non-matching trace events between the event log and the process model to reach an optimal alignment. As for the denominator, $C_{L_{Non-SynchonousMoves}}$ is the total cost required for aligning all activities in the event log with the process model, assuming the worst-case scenario where no perfect matches exist. Therefore, fitness values closer or equal to 0, indicate underfitting meaning that the process model fails to represent the execution log sequences. Conversely, fitness values closer or equal to 1, indicate that the process model can properly reproduce the execution behavior contained in the event log.

$$Precision = 1 - \left(\frac{\sum_{state\ \boldsymbol{i}}^{state\ \boldsymbol{n}}\left(Visit\ Counts * \left(\frac{OutgoingEdgeCount - FiredEdgeCount}{OutgoingEdgeCount}\right)\right)}{\sum_{State\ \boldsymbol{i}}^{State\ \boldsymbol{n}} Visit\ Counts}\right) \quad \text{(Eq. 8)}$$

Where $Precision = \frac{\sum(Observed\ Behavior)_L}{\sum(Allowed\ Behavior)_M}$, and it can be represented in terms of traces, events, or scaping edges as in (Eq. 8). Therefore, precision values close to or equal to 0 indicate overgeneralization or poor precision due to the high fraction of allowed edges in the process model that are not observed in the log. Oppositely, if there are no escaping edges or no additional behavior is allowed by the discovered process model, then the precision is high, and near to 1.

5.4.2.2. Module 5 (M5): Process performance assessment

Once the actual process behavior has been discovered and validated, the derived process model(s) serves as input for automated PHM, which involves detecting and diagnosing real-world process performance issues in E2E business operations by dynamically assessing their performance. In this regard, this study implements the proposed LPMM methodology to perform the performance assessment of the RFI management process. This hybrid methodology not only helps reveal process inefficiencies but also provides data-driven explainability of why they might be occurring, thereby generating valuable and actionable process performance insights.

For the quantitative process performance assessment, this study focuses on four main process perspectives of PHM. Table 5-1 provides an overview of the first two perspectives. First, the control-flow perspective where a dominant process model for the case study is automatically discovered from reconstructed evet logs, then it is validated using (Eq. 7) and (Eq. 8). Based on this derived 'as-happened' process model, main process deviations, frequency-based bottlenecks, and rework loops are identified. Additionally, a cross-case process variant analysis is carried out to detect and compare performance patterns of actual process executions together with a text mining n-gram and Term Frequency – Inverse Document Frequency (TF – IDF) analysis for the unstructured data content. Second, the time perspective comprises general metrics such as: (i) inter-arrival, on time and late response rates; (ii) Lean-based metrics into how specific process behavior patterns impact activity cycle times, lead time, and Work in Progress (WIP); (iii) a comparative process mining implementation is employed to analyze the Weighted Average Process Efficiency (WAPE) of the executed business process (i.e., time-based efficiency) across projects using (Eq. 9), which was derived from value-added process time Lean-based concept [225, 226]; and (iv) process drift and batch detection using performance spectrum miner (PSM) [224, 238]. For instance, (Eq. 9) provides an example of computing WAPE for RFI management; however, the formulation can be generalized by replacing the RFI entity with other process entities, such as change orders or progress payments.

$$WAPE_{Project\ i}\ (\%) = \frac{\sum_{v=1}^{v=n}\left(\frac{VAT_{RFIs\ iv}}{Lead\ Time_{RFIs\ iv}}\right)}{\sum No.of\ RFI\ Variants} * 100 \quad \text{(Eq. 9)}$$

Where $\sum_{v=1}^{v=n}\left(\frac{VAT_{RFIs\ iv}}{Lead\ Time_{RFIs\ iv}}\right)$ represents the sum of time-based ratios for each process variant in project, from the first variant (v = 1) to the last variant (v = n); $VAT_{RFIs\ iv}$ = Value Added Time (i.e., Cycle Time) of all RFIs in variant v; $Lead\ Time_{RFIs\ iv}$ = Lead time of all RFIs in variant v; and $\sum No.of\ RFI\ Variants$ = total number of process variants in project i.

Table 5-2 presents an overview of the two remaining perspectives examined in the quantitative process analysis. The cost-oriented perspective is enabled by enriching the event log with cost-related attributes, as described in Figure 5-5. In this study, the cost perspective includes general descriptive statistics and specific key process performance metrics such as the total cumulative process cost, the average cost of the discovered dominant process, identification of costliest process variants, and the computation of the Weighted Average Cost Efficiency (WACE) as per the deducted (Eq. 10). This equation provides an example of computing WACE for RFI management; however, the formulation can be generalized by replacing the RFI entity with other process entities, when analyzing other business processes.

$$WACE_{Project\ i}\ (\%) = \frac{\sum_{v=1}^{v=n}\left(\frac{Cost_{RFIs\ iv\ within\ Threshold}}{Cost_{RFIs\ iv}}\right)}{\sum No.of\ RFI\ Variants} * 100 \qquad \text{(Eq. 10)}$$

Where $\sum_{v=1}^{v=n}\left(\frac{Cost_{RFIs\ iv\ within\ Threshold}}{Cost_{RFIs\ iv}}\right)$ represents the sum of cost-based ratios for each process variant in project, from the first variant (v = 1) to the last variant (v = n); $Cost_{RFIs\ iv\ within\ Threshold}$= the total cost of all RFIs within variant v of project i that fall within the defined threshold; $Cost_{RFIs\ iv}$ = denotes the total cost of all RFIs within variant v; and $\sum No.of\ RFI\ Variants$ = total number of process variants in project i.

The fourth perspective is the organizational one, which examines team dynamics among resource participants involved in executing the RFI process. Within this perspective, key social networks are identified, including Handover-of-Work, Subcontracting, and Collaboration networks. For each network, various SNA performance metrics [90] are analyzed as described in Table 5-2.

By identifying actual process inefficiencies or operational performance pain points across multiple process perspectives through the proposed LPMM methodology, construction organizations can leverage their project operations data to extract actionable performance insights. This enhances process transparency and visibility into performance status, supports fact-based decision-making by key project stakeholders, and enables targeted data-driven process improvement strategies.

5.4.3. Process-driven decision-making – Level 5 (L5)

The upper level (L5) of the process-oriented architectural layer of the D2P method, depicted in in Figure 5-1, comprises the two top-level modules of the LPMM framework, which aim to drive process intelligence by supporting data-driven decision-making based on the process-driven insights discovered during the performance assessment and diagnostics phase as described in previous sections. These modules constitute cumulative components of the proposed process-aware framework for the continuous monitoring and process management, as they promote evidence-based decision-making and support automated, continuous PHM. Collectively, these capabilities facilitate the systematic identification of critical process performance deficiencies, thereby informing targeted improvement strategies to streamline critical business operations and control process variability. These capabilities help reduce risks associated with this variability.

5.4.3.1. Module 6 (M6): Process improvement opportunities

Module 6 (M6) focuses on identifying timely, targeted interventions for process improvement and variant-level process control, informed by data-driven decision points extracted from actual process executions and grounded in diagnostic analysis. The resulting improvement opportunities comprise data-driven best practices derived from observed performance patterns and are formulated as action- and target-oriented intervention recommendations based on the extracted performance insights [72]. This study is grounded in the premise that a data-driven diagnostic analytical phase to evaluate and monitor "as-happened" process performance must precede any business process automation, simulation, or process re-engineering initiative to ensure that interventions systematically address the most critical root causes of process inefficiencies and sustain robust variant-level process control while preventing wasteful and premature automation efforts.

5.4.3.2. Module 7 (M7): Continuous Process Performance Monitoring & Dynamic Reporting

Module 7 (M7) of the proposed process-oriented framework focuses on continuous PHM through customized reporting implemented via dynamic process performance boards, leveraging process mining–based insights derived from the quantitative process assessment enabled by the D2P method. These reporting mechanisms provide sustained visibility into E2E process performance, thereby enhancing transparency and enabling continuous monitoring and process auditing. Section 6.2.3.2 presents a real-world implementation of such dynamic process performance boards. In addition to the key performance metrics defined in Table 5-1 and Table **5-2**, key PPIs can be defined by AEC/FM organizations for instance from AACE Total Cost Management Framework [239], project contracts [240], construction schedules [138], execution plans [139], or expert inputs[46] the latter (PPIs) differ from metrics in that they define operational performance targets to be monitored as part of the PHM, for example benchmarking targets for RFI resolution times [241].

CHAPTER 6. METHOD IMPLEMENTATION AND RESULTS: RFI MANAGEMENT[5]

This chapter focuses on the application of the proposed LPMM framework through a real-world case study involving an initial dataset of more than 1,400 construction projects comprising over 8,570 RFIs managed by a large general contractor (GC) in North America. The analyzed projects span four distinct lifecycle phases: bidding/budgeting, pre-construction, construction, and post-construction. The study consists of assessing the operational performance of the RFI management process through the implementation of the proposed LPMM method across multiple project types. The performance evaluation focuses on four key perspectives: the control-flow and time, described in Table 5-1, and the cost and organizational, presented in Table 5-2. These tables also outline the performance aspects and questions to be investigated in this case study, along with the defined key PPIs to be measured. While all four perspectives are examined, the primary focus is on process efficiency, particularly in terms of time and cost. The case study has three primary objectives: (i) to apply the LPMM framework in real-world settings; (ii) to uncover latent inefficiencies and their potential root causes by quantitatively assessing the actual execution performance of the RFI process in an automated manner; and (iii) to identify process improvement opportunities derived from the quantitative assessment.

Although the study includes descriptive analytics and derived prescriptive recommendations (i.e., derived from diagnostics), the performance assessment primarily centers on diagnostic analytics grounded in process mining techniques and Lean-based metrics. Diagnostic analytics play a critical role in PHM, as they can reveal "as-happened" process performance issues of RFI workflows as executed by project stakeholders. The process-aware insights extracted from

[5] This chapter is based upon:
A. J. Martinez Lagunas and M. Nik-Bakht. 2025. "Automated RFI End-to-End Process Performance Assessment and Monitoring: A Digital Mining-based Method to Enhance Construction Operations Management". Status: Submitted – Under Review

this analysis are essential for enabling and enhancing advanced predictive and prescriptive analytics. Most predictive models lack explainability and transparency concerning the underlying process dynamics, specifically in understanding how the process is actually executed and in identifying the most likely reasons behind the predicted suboptimal performance. To address this gap, the study extracts actual performance behavior patterns and identifies their underlying root causes. This case study also aims to provide insights for construction organizations seeking to transform their data-oriented systems into process-aware architectures. It presents a general approach for deriving actionable performance insights from project operations data related to RFI management, identifying targeted opportunities for process improvement and supporting data-driven decision-making. The following sections are organized in alignment with the methodological framework introduced in Chapter 5. , with a specific focus on presenting the RFI management process case study and its results. These subsections provide further details on the projects examined in this study, the types of data related to the RFI process, the implementation steps of the PHM framework, and the findings derived from the automated process performance assessment.

6.1. LPMM Framework - Layer I: Data-oriented architecture

6.1.1. Data-oriented systems, project roles, and raw data – Level 1 (L1)

The GC largely relies on Procore [199] and SQL Server [218] databases to securely store, query, share, and manage its business operations data associated with its construction project portfolio. The SQL Server database snapshot contains a ‘static data view’ of the GC’s project portfolio with real-world operations data. The static data view was extracted by the GC in August of 2023. The ‘procore.projects’ data table contained 1,402 construction projects, mostly located in the United States, being managed by the GC using Procore, as of the date of data extraction. Although this static SQL database snapshot included some historical project data, it lacked the required dynamic, stateful, process-aware data, specifically RFI process lifecycle traceability. Accordingly, two key actions were undertaken to address this challenge. The first action was to develop a Procore data connection application on the Procore Developers Portal [206, 235], enabling the dynamic extraction of specific stateful process data for key projects directly from the GC’s Procore Instance using HTTP REST API calls method [241]. This type of extraction was performed whenever the required stateful data attributes were unavailable in the SQL database snapshot. The second action involved extracting anonymized insights into the number of user actions performed by construction stakeholders on the GC's Procore instance using a task mining approach [125]. The task mining scope in this study consisted of extracting project stakeholders' user activity from Procore and analyzing it to gain a general understanding of the most active projects, key stakeholder roles, the most frequently used Procore modules and tools, and the overall significance of the RFI management process within those projects. It is worth noting that the outcome of these two actions provided dynamic insights, in contrast to the static views offered by the SQL data snapshot only. Therefore, the enriched data collection for this study comes from both data sources: Procore and SQL.

Figure 6-1 illustrates the results of the GC's task mining analysis based on Procore usage data, revealing that within the Procore Project Management module, the top three most-used tools at the time of the analysis were Drawings, Submittals, and RFIs. Although the drawing and submittal approval processes are highly important, this study focuses primarily on the RFI management process. It is worth noting that drawings and submittals are often referenced in RFIs by involved parties or even attached to them; however, the specific analysis of those processes falls outside the scope of this study. This presents an opportunity for scholars and practitioners to explore these processes in future research, as highlighted in [26]. The most active stakeholder roles and the most active projects are determined by the number of user actions performed within each project and tool as shown in Figure 6-1. The usage analysis included activity from office-based teams, who tend to use Procore Web, and field-based teams, who primarily use Procore Mobile.

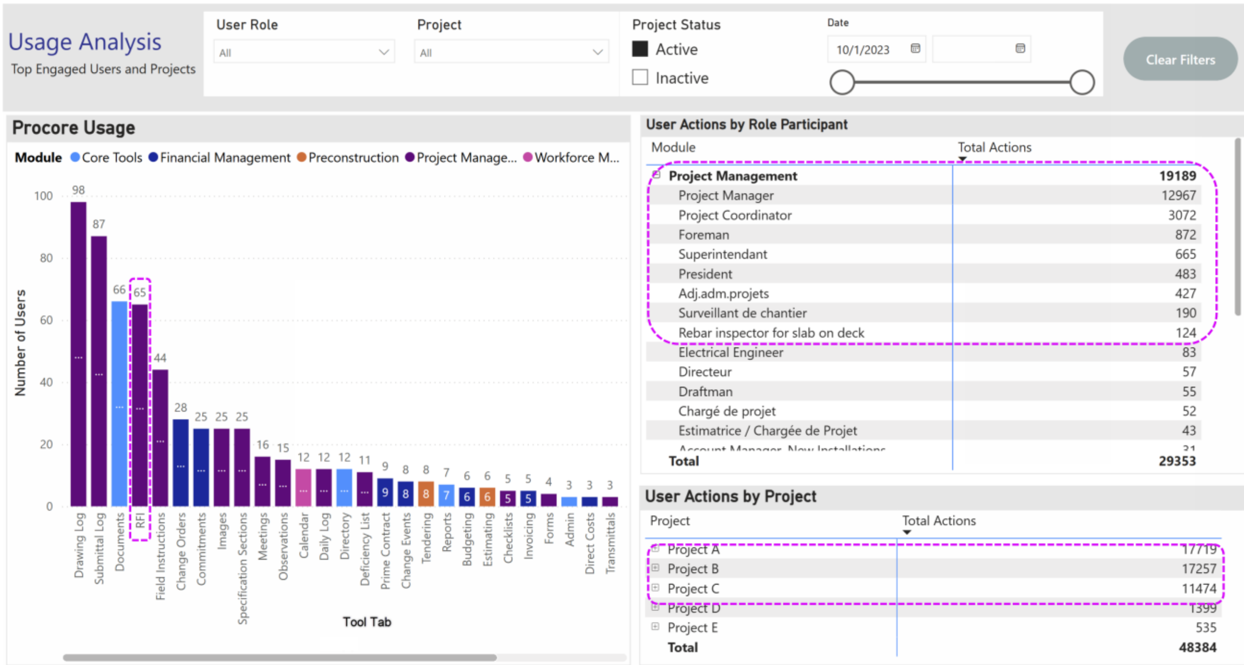


Figure 6-1. Sample users' usage analysis on Procore CDE by project and stakeholder role – Power BI.

This stage also involved identifying relevant raw Procore data related to the RFI management process, including, among others: RFI identifiers, lifecycle statuses, timestamps, user identifiers, locations, disciplines, drawing numbers, and free-text fields such as questions, subjects, and responses. As summarized in Table 6-1, the raw RFI-related data tables identified in the GC's SQL database include: Companies, Projects, Workers, Project Roles, RFIs, RFI Questions, RFI Answers, and RFI Drawings.. Table 6-1 presents the number of records and attributes of each dataset, along with their identifiers and linking attributes, namely the primary and foreign keys.

Table 6-1. GC's SQL database snapshot: raw data overview.

Data Table	No. of Datapoints	No. of Attributes	Primary Key	Foreign Keys
Companies	11,362	71	CompanyID	EmployeeID
Projects	1,402	65	ProjectID	CompanyID
Workers	22,281	45	WorkerID	CompanyID
Project Roles	15,895	15	RoleID	WorkerID; ProjectID
RFIs	8,578	71	RFI_ID	CompanyID; ProjectID; WorkerID
RFI Questions	8,578	18	RFIQuestionID	RFI_ID; ProjectID
RFI Answers	11,236	20	AnswerID	RFI_ID; RFIQuestionID; ProjectID
RFI Drawings	2,125	14	DrawingID	RFI_ID; ProjectID

6.1.2. Data cleaning, data model, and project information – Level 2 (L2)

At L2, the initial data cleaning process involved the following steps: first, any projects without logged RFI data in the available data sources were filtered out from the initial set of 1,402 projects; second, Procore sandbox test projects and projects with estimated contract values ≤ $1 (USD)were excluded, as they were created solely for testing and do not represent real projects. This resulted in 215 projects with a total of 8,349 RFIs, which were subsequently narrowed down according to the criteria outlined in the sections below for this study.

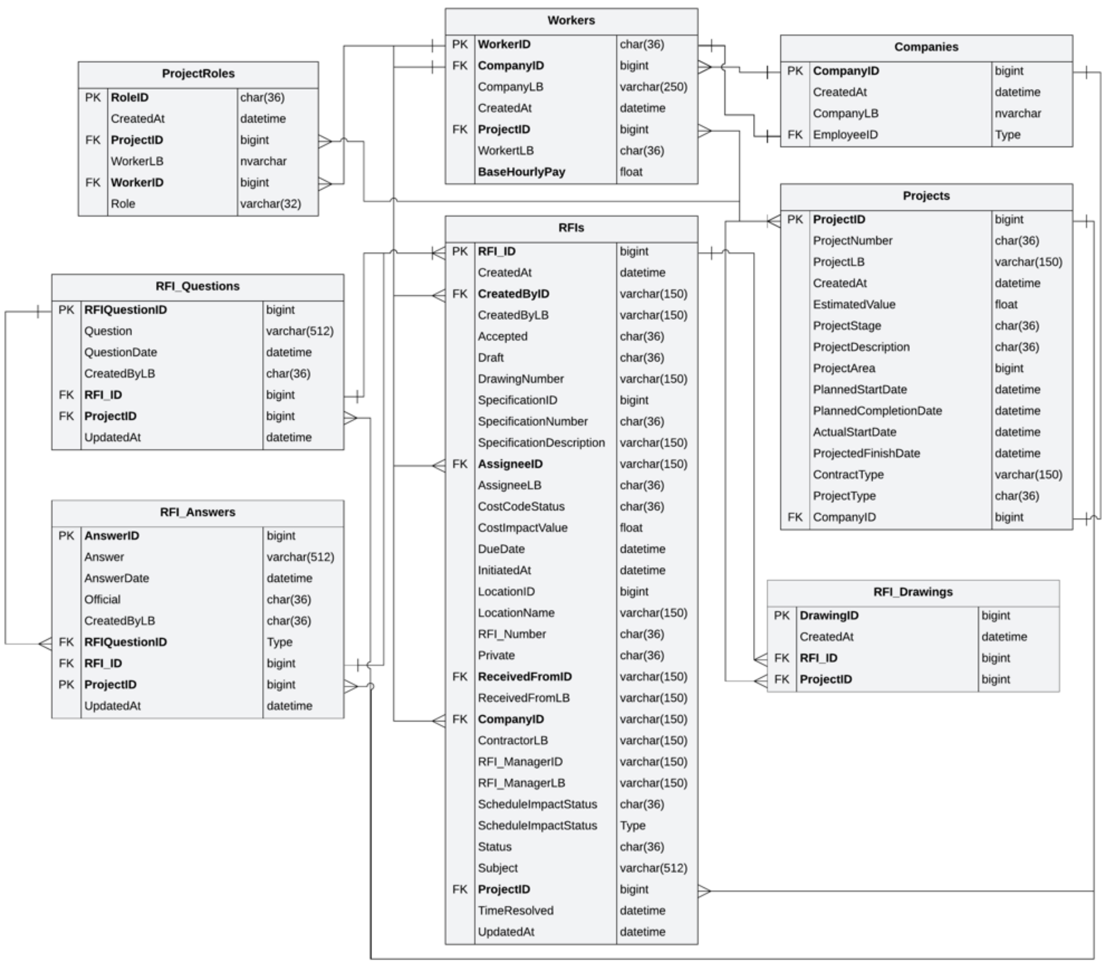


Figure 6-2. Reconstructed Entity-Relationship Diagram (ERD) of RFI management SQL data tables.

Based on the identified primary and foreign keys of each data table described in Table 6-1, a CDM related to the RFIs across these projects was generated and represented as an ERD, shown in Figure 6-2. Each table represents a data model entity, and these entities are interrelated through one of the following relationship types: many-to-many, one-to-one, one-to-many, or many-to-one. For example, multiple 'Project Roles' can participate in the RFI process of various 'Projects', a 'Project' can have multiple 'RFIs', each 'RFI' is associated with one 'RFI' question, an 'RFI Question' can have multiple 'RFI Answers', and each 'RFI' can refer to multiple 'Drawings'. Furthermore, as shown in Figure 6-2, not all raw attributes were included in the ERD; all null or redundant data attributes, as well as any irrelevant metadata, were excluded.

6.2. Layer II: Process-aware architecture

6.2.1. Foundational Process Data Models (FPDMs) – Level 3 (L3)

As described in Section 5.4.1, the generation of FPDMs is an essential step for AEC/FM organizations seeking to transform their current data-oriented architectures (i.e., Layer I in Figure 5-1) into process-aware structures (i.e., Layer II in Figure 5-1). This step L3 consists of three main subtasks: (i) understanding the process of interest; (ii) designing the RFI management process mining use case, and (iii) producing the required event logs in accordance with the XES standard [108]. The produced event logs are the backbone of the foundational process-oriented data models.

6.2.1.1. Module 1 (M1): Process and project selection, process understanding

Process Selection – This study focuses on analyzing the RFI management process not only because it is identified in the literature as a core process [26, 242], but also due to its observed significance across all construction projects, regardless of the adopted delivery system, and its relevance to the projects managed by our industry partner. In this regard, the starting point is to gain a general understanding of the process, either from recommended best practices in the existing literature [24, 243] or from 'as-planned' process models documented by construction organizations. The expected RFI process model is often tailored to meet specific company's organizational requirements, project stakeholder roles, project settings, delivery methods, and contractual agreements.

Project Selection – Apart from identifying the underlying CDM (see Figure 6-2), another important aspect is determining how many RFIs are managed in those projects. The volume of RFIs managed in each project is a key feature from a process-oriented standpoint, as each RFI represents a specific workflow instance involving several interrelated activities often executed in different ways by multiple project stakeholders. These instances may include various decision points, required approvals, and other elements depending on the size, type, and complexity of the construction project.

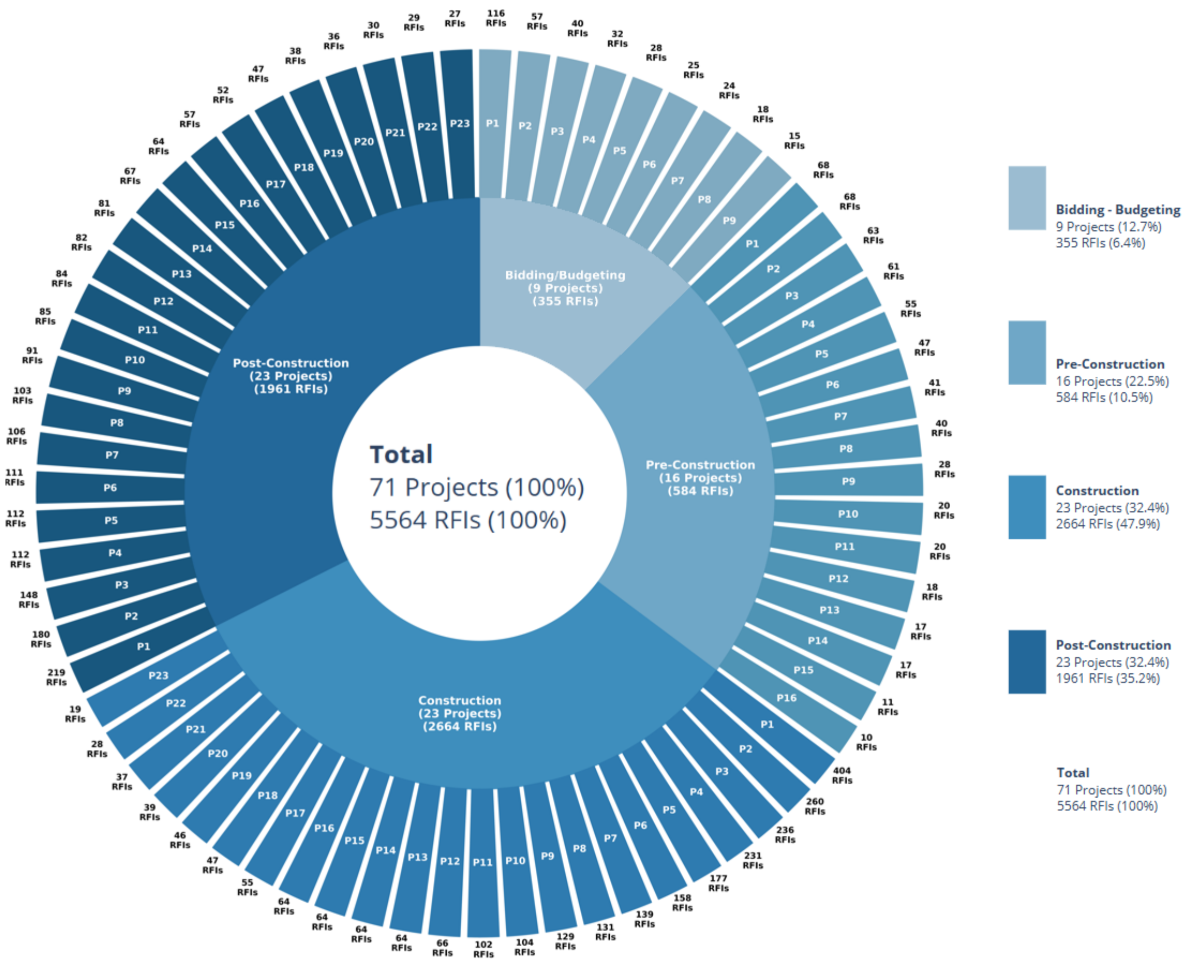


Figure 6-3. Breakdown of selected RFI project data for process-oriented analysis (see interactive chart here).

Taking these aspects into account, the following logic was applied to select the projects and their corresponding RFIs for quantitative process assessment using the proposed LPMM method, as presented in the subsequent sections: (i) all projects were grouped by lifecycle phase, project type, and contract type; (ii) projects were ranked based on the highest volume of RFIs and the maximum estimated project values, from which the top two were selected due to their greater significance. In cases where only one project existed for a specific project type, that top project was also selected; and (iii) based on the statistical quartiles of RFI counts within each project group, it was observed that some projects selected in the previous step still exhibited relatively low RFI volumes compared with the minimum RFI count in the highest quartile. To ensure sufficient process activity for analysis, an additional filtering threshold was therefore applied, excluding projects with fewer than 10 RFIs. After applying these filtering criteria, the final dataset comprised 71 projects with a total of 5,564 RFIs. Figure 6-3 presents an overview of these projects by project phase, along with the corresponding RFI count for each project. This final cleaned dataset included project types such as industrial, residential, healthcare, institutional, and commercial projects.

Process Understianding – Extended Literature Review on RFI management: Given its high importance, the existing literature on RFI management and its performance assessment in construction projects, as reported by several scholars and industry professionals, was meticulously reviewed and classified into three major streams based on the primary analytical approach adopted by the studies: (i) text mining-based studies; (ii) data-oriented studies; and (iii) process-oriented studies as summarized in Table 6-2.

Aiming to streamline document management workflows, Das et al. (2022) designed a blockchain-based framework that leverages (i) smart contract technology to customize approval workflow logics based on predefined client requirements; (ii) blockchain ledgers to maintain a traceable record of document lifecycle state changes; and (iii) document management tree-like structures to guarantee version control. The authors then developed a proof-of-concept (POC) prototype for the RFI process, enabling automated management features with enhanced process control [244]. In the same vein, Erri Pradeep et al. (2021) developed a blockchain-based process model prototype to support the management of design-related RFIs by keeping a record of all historical RFI transaction updates [245]. Even though both studies [244, 245] propose advanced features to enhance and automate the traceability of the RFI process, there remains a need to identify and diagnose actual process inefficiencies to enable the implementation of laser-focused automation initiatives.

Another aspect examined by various research scholars is process interoperability, which aims to enable seamless and standardized process implementation across multiple projects and organizations. Golzarpoor et al. (2017) developed the so-called “Industry Foundation Process (IFP) ontology” to facilitate automated process conformance checking of core business processes, including the RFI process, by applying object-oriented programming rules, and to foster process interoperability [24, 53, 169]. In line with this effort, Zhu and Augenbroe (2006) proposed a conceptual model using Unified Modeling Language (UML) notation to enrich existing cloud-based, data-oriented Application Programming Interfaces (APIs) with process-oriented functionalities to support intra- and inter-organizational information process management. The model comprises four main interrelated process components: tasks, cases or documents, organizational resource teams (i.e., task performers), and the decision criteria associated with those tasks [25]. Regardless of the contributions of these studies, they primarily adopted a top-down (normative or prescribed) approach rather than considering a bottom-up (data-driven) approach to continuously derive process-oriented performance behavior from actual operations data [26].

When it comes to analyzing the process performance of RFI management, only a few efforts have been reported in the literature. Mohamed et al. (1999) developed a dynamic simulation model to estimate the time and cost impacts of the RFI process, as well as the potential savings generated by executing it through cloud-based solutions such as construction Project Management Information Systems (PMIS). The RFI process was modeled as a series of business tasks transferred from one process actor (i.e., task performer) to another. The study reported that each RFI requires a mean processing time of 17 working hours, incurs an estimated cost of about 1,380 Australian dollars, and could achieve potential time savings of up to 40% through PMIS implementation [246].

Similarly, adopting a simulation-based approach, Papajohn and El Asmar (2020) used discrete event simulation (DES) to model the RFI process as a stochastic series of events. Their study investigated the RFI response time for 2,000 RFIs from 17 highway projects under Design-Build (DB), Design-Bid-Build (DBB), and Construction Management/General Contractor (CM/GC) systems, assigning each RFI a specific urgency level based on practitioner survey responses [247]. Both studies [246, 247] applied Monte Carlo random sampling to account for the uncertain nature of RFIs, particularly for tasks whose durations vary significantly depending on their complexity. Notably, in contrast to the findings of Kim et al. (2022) [248], which also analyzed civil projects, the study in [247] found that DB projects had a greater number of RFIs and longer response times than those delivered under DBB or CM/GC. The study attributed these findings to intrinsic characteristics of DB projects, such as greater uncertainty, complexity, and urgency. However, both simulation-based studies [246, 247] relied on subjective input from construction practitioners to model the process, and even though [247] incorporated actual logged durations for certain RFI tasks, the study acknowledged that aspects such as team dynamics of the actors performing those tasks and actual E2E process variability were not investigated.

Both described simulation-based studies [246, 247] offer a forward-looking approach to assess what-if performance scenarios of the RFI process, nonetheless they do not consider the importance of diagnosing the actual E2E process performance. In other words, they overlook the use of modern data-driven approaches such as process mining to enrich simulation model inputs by extracting actual process performance behavior, which in turn can help improve simulation results and process efficiency [26].

As noted by several scholars and industry reports, modern ICT developments play a key role in enhancing the efficiency of the RFI management process and mitigating potential risks, by preventing the abuse or overuse of RFIs and supporting timely responses [242, 249]. This process is an essential communication tool in construction projects regardless of the chosen delivery method [227, 246, 250]. In this context, although several cloud-based construction management platforms are available on the market, Aconex [211] and Procore [199] are among the most widely adopted by construction organizations to aid in managing construction information processes, with Procore having a larger presence in North America [199, 251]. However, at present, most of these construction platforms, although continuously evolving, rarely enable dynamic assessment and continuous performance monitoring of actual E2E processes [26]. In this regard, various recent studies have emphasized the opportunity of leveraging novel technologies and methodologies such as Building Information Modeling (BIM), Natural Language Processing (NLP), and Large Language Models (LLMs), blockchain, and process mining to enhance, automate, and control the information management processes [26, 72, 232, 242, 244, 252, 253].

Table 6-2. Synthesis of previous studies in relation to the proposed LPMM Framework

Authors	Ref	Year	Location	RFI	CPA	TM	DO	PRO	PER	DEA	DIA	PDA	PSA	PRM	E2E	AS	AT	TB	T	C	CF	O
Afzal et al.	[254]	2023	Australia	✓		✓				✓												
Yilmaz & Ergen	[255]	2024	Greece	✓		✓	✓			✓												
Hanna et al.	[243]	2012	USA	✓	✓		✓		✓	✓									✓			
Bhat	[256]	2017	Canada	✓	✓		✓		✓	✓									✓			
Ozogul & Ergen	[257]	2024	Greece	✓		✓	✓			✓												
Panahi et al.	[258]	2023	USA	✓		✓	✓					✓										
Mao et al.	[259]	2007	USA	✓			✓			✓												
Kim et al.	[248]	2022	USA	✓	✓		✓		✓	✓										✓		
Adamtey	[260]	2021	USA		✓		✓		✓	✓									✓	✓		
Aibinu et al.	[261]	2020	Australia	✓	✓		✓												✓			✓
Ibrahim et al.	[262]	2020	USA	✓			✓			✓									✓	✓		
Das et al.	[244]	2022	China	✓				✓		✓					✓	✓						
Pradeep et al.	[245]	2021	New Zealand	✓				✓		✓												
Golzarpoor et al.	[24, 169]	2016 2018	Canada	✓				✓		✓					✓	✓	✓				✓	
Papajohn & El Asmar	[247]	2020	USA	✓	✓			✓	✓	✓		✓				✓		✓	✓			
Mohamed et al.	[246]	1999	England	✓			✓	✓	✓	✓		✓				✓		✓	✓	✓		✓
Kouhestani & Nik-Bakht	[253]	2020	Canada				✓	✓	✓	✓	✓			✓	✓	✓	✓		✓		✓	✓
Pan and Zhang	[90]	2021	SGP				✓	✓	✓	✓	✓			✓		✓	✓		✓		✓	✓
Wang et al.	[263]	2024	China		✓		✓	✓	✓	✓	✓			✓	✓	✓	✓		✓		✓	
Martinez & Nik-Bakht (This Study)		2025	Canada	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓

(RFI) RFI related; (CPA) comparative project analysis; (TM) text mining related; (DO) data-oriented; (PRO) process-oriented; (PER) performance-based; (DEA) descriptive analytics; (DIA) diagnostic analytics; (PDA) predictive analytics; (PSA) prescriptive analytics; (PRM) process mining-based; (E2E) E2E process model; (AS) as-planned model; (AT) actual model; (TB) to-be model; (T) time perspective; (C) cost perspective; (CF) control-flow perspective; (O) organizational perspective; (USA) United States of America; and (SGP) Singapore.

With regard to other early efforts on the implementation of process mining in the construction domain [26, 35], Pan and Zhang (2021) applied fuzzy and inductive process mining algorithms to discover the actual process flow related to the execution of architectural and structural works and analyzed it from an organizational perspective, that is, for instance, by examining workers' workloads and bottlenecks, including the evaluation of team dynamics performance through various Social Network Analysis (SNA) metrics [90]. Moreover, Wang et al. (2024) proposed enhancing BIM-based digital twins of highway projects by integrating process-oriented insights derived from process mining discovery algorithms. They quantitatively analyzed the actual performance of different civil and structural work process variants for each construction project in terms of frequencies and time durations and compared the 'as-planned' against the actual or 'as-happened' performance of these construction processes[263].

As a result of an exhaustive literature review [26], including the studies discussed in this section, it is worth noting that no studies were found in the existing body of knowledge that provide a systematic approach tailored to the construction domain for implementing process mining technology enriched with lean-based metrics to automatically extract and quantitatively assess the E2E process performance. This gap, particularly the lack of a domain-specific quantitative method to support the evaluation of business operations' agility in construction projects. This is what this study aims to address. Table 6-2 and Table S10 present an overview of the existing related literature alongside this study's main characteristics, highlighting the additional aspects and components addressed by this study that have not been covered in previous research.

Process Understanding – RFI Process Description: The operational significance of the RFI process across all project lifecycle phases for successful project delivery has been reported in prior studies [26, 46, 242], In this context, this step consisted of establishing a reference understanding of the RFI process based on recommended best practices documented in the literature [24, 243] and on an "as-planned" RFI process model adopted by large general contractors, informed by aggregated expert input [46]. This reference model (see **Figure 6-4**) mapped the required inter-organizational RFI management procedures involving the GC, including stakeholder role interactions across organizations, project settings, information delivery mechanisms, and contractual arrangements.

Figure 6-4 presents the "as-planned" reference RFI management process model typically implemented by large contractors in North America [24, 46, 243]. In this model, RFIs are typically initiated by a Specialty Contractor (SC) or by the General Contractor (GC). The initiator formulates the RFI using a standardized template and records it in the CDE, where it may first be saved as a draft for internal review prior to formal submission. Depending on the initiator, the RFI is submitted from the SC to the GC or from the GC to the Owner Representative (OR), typically an architect or consulting specialist. Submission is automatically confirmed through system notifications. Upon receipt, the GC or OR verifies the RFI's completeness, clarity, and applicability. Valid RFIs are assigned by the RFI manager to appropriate reviewers and distributed to relevant stakeholders according to discipline and subject. When escalation is unnecessary, the GC or OR reviews and responds directly to the SC or GC. Otherwise, the RFI is reassigned to the appropriate reviewer before a response is issued. Accepted responses typically lead to RFI closure, although in some cases they may trigger the initiation of a change order request process. If the response is rejected or unclear, the initiator reformulates and resubmits the RFI, restarting the review cycle. RFIs closed unintentionally may be reopened to resume the process if required.

In practice, multiple process variations may occur relative to the reference RFI process described above (see **Figure 6-4**). Such deviations can significantly affect process performance, disrupt project operations, and potentially lead to schedule delays and cost overruns. It is also important to note that RFIs may arise during any phase of the construction project lifecycle, as illustrated in Figure 6-3. Accordingly, quantitative assessment of RFI process performance is essential to enable timely process control interventions when suboptimal performance is detected.

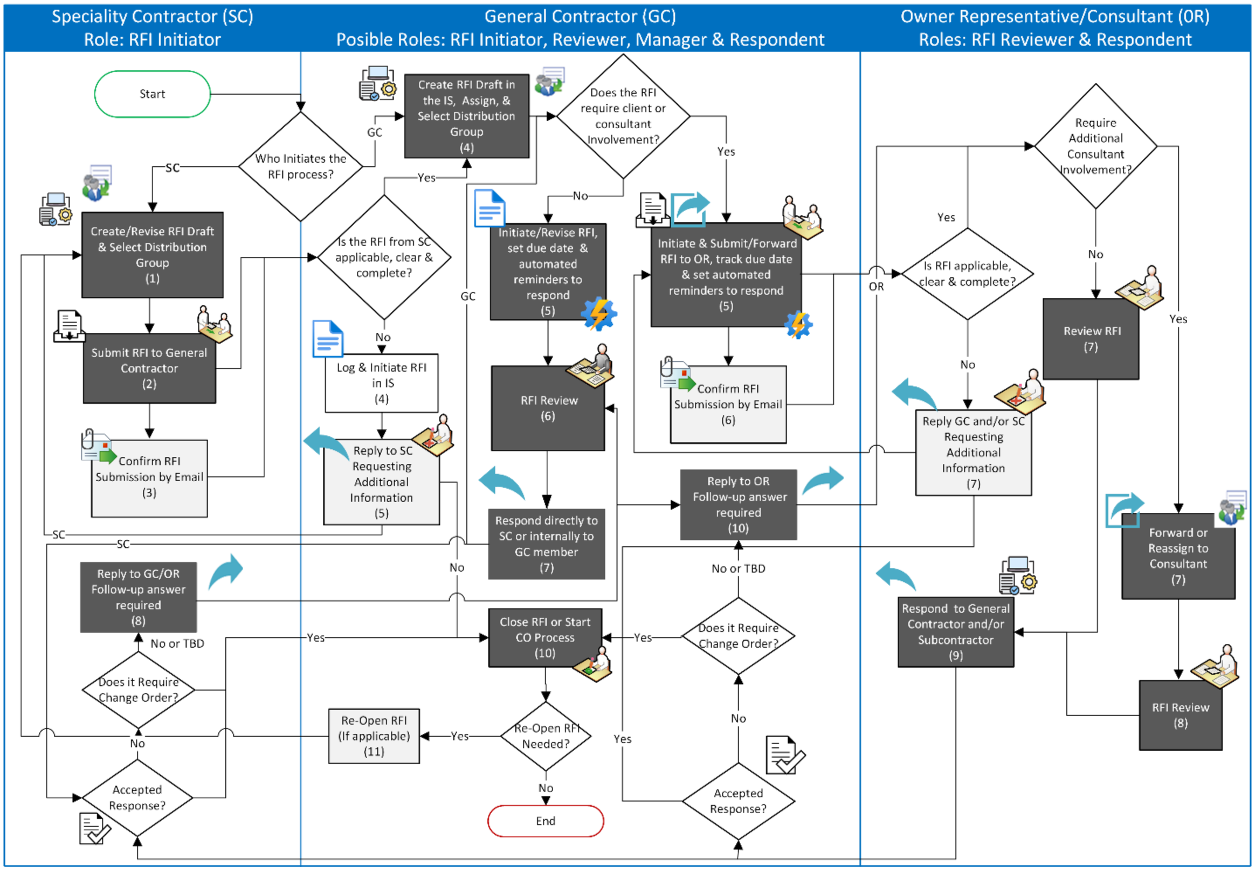


Figure 6-4. "As-planned" RFI management process: baseline workflow activities highlighted in dark gray.

6.2.1.2. Module 2 (M2): Process Mining use case development and scope definition

This section presents the development of the RFI management process mining use case in line with section 5.4.1.2. As depicted in Figure 6-5, the use case design consists of three major components: (i) the construction PMIS module supporting the execution of the RFI management process; (ii) the ETL processing component corresponding to data extraction, transformation, integration, and event log generation; and (iii) the process mining component.

The first component provides a high-level view on the construction organization's technological landscape related to the information management of RFI operations data of the GC's project portfolio; in this case the GC uses Procore as the main PMIS. More specifically, Figure 6-5 illustrates the three main Procore submodules that support RFI management: the Project Management submodule, and two interrelated submodules, Financial Management, which integrates contractual data from Oracle NetSuite [200], and Workforce Management, which utilizes the Field Productivity application via Procore Mobile. Additionally, the specific Procore tools used to execute the RFI process within these submodules are identified. Finally, the roles of project stakeholders interacting with these tools and module components are also mapped. All these aspects represent the identified HMI related to the RFI process, primarily derived from the task mining analysis of Procore user activity presented in Figure 6-1.

The second and third components, illustrated in Figure 6-5, form the backbone of the LPMM method proposed in this study; therefore, they are discussed in more detail in the following subsections. However, all in all, the second component primarily focuses on event log generation in accordance with the XES international standard, while the third component centers on leveraging process mining techniques for automated reconstruction or extraction of actual RFI process executions from event logs, PHM from multiple analytical perspectives, and dynamic performance reporting based on quantitative assessment.

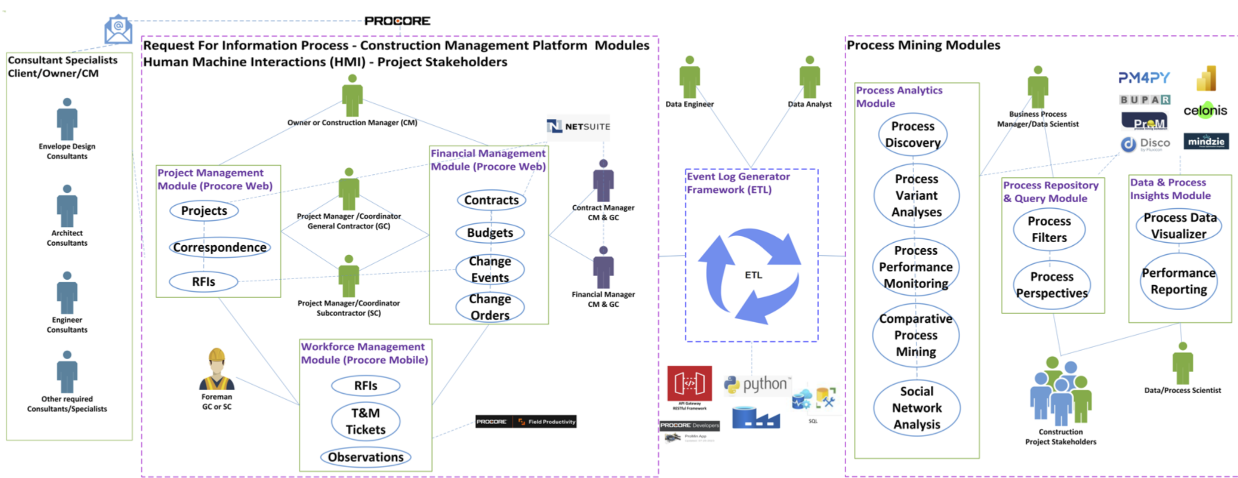


Figure 6-5. RFI management process mining use case

6.2.1.3. Module 3 (M3): Process event log generator

The generation of process event logs is not a trivial task and it normally requires to transform raw systems' data into a well-defined process-aware structures [232, 233]. This task can take up to 80% of the process mining implementation efforts [76, 103]. To ensure robust implementation of the proposed LPMM method, it is necessary to produce well-structured process event logs, meaning that these event logs must be machine-readable with proper syntax and semantics. Therefore, most construction organizations lack a systematic approach to evolve their data-driven technology landscape into an architecture that integrates core business process awareness, thereby enabling the automated execution of this task.

In this case, RFI management operations data was extracted from the GC's Procore web instance using the developed extractor Figure 5-3, as described in Section 5.4.1.3. As shown in Figure 5-3, this algorithm uses HTTP REST API method and the developed Procore data connection app to enable the RFI data extraction.

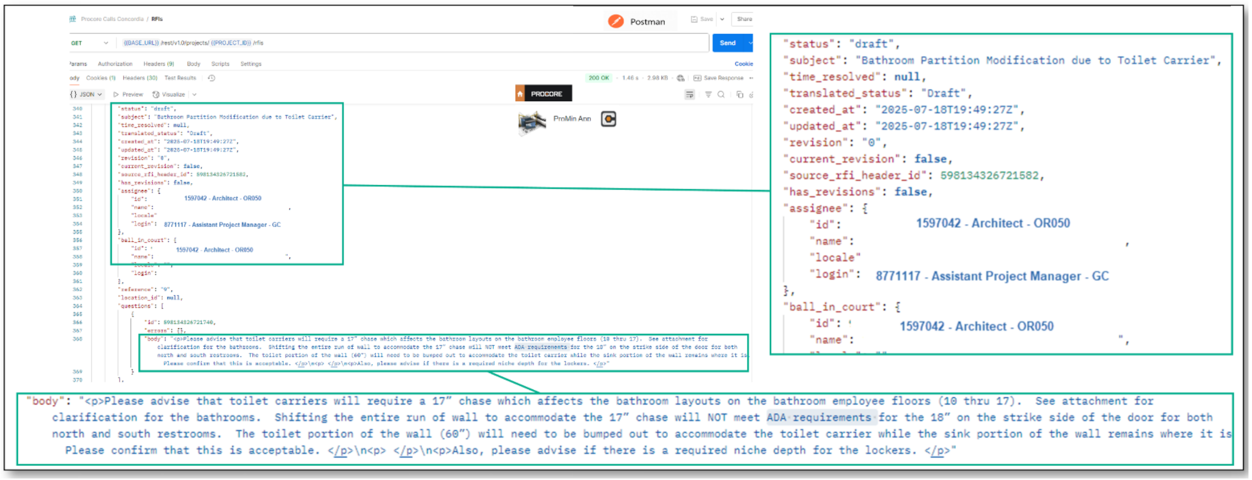


Figure 6-6. API-based RFI data extraction from Procore construction management platform using Postman.

To demonstrate the outcome of Figure 5-3, Figure 6-6 displays an actual example of RFI data extraction for one of the GC's major commercial construction projects. It is worth noting that, in order to feed the next algorithm that transforms raw system data into process event logs, specific data attributes are required. These include, among others: the RFI instance identifier, process lifecycle timestamps, RFI status, assignees, reviewers, project stakeholders' roles, RFI subject, and questions. Such attributes are essential for constructing robust event logs, as they provide the necessary contextual information about the HMI. For instance, they indicate what RFI-related actions are being executed in the system, when they occur, what information the RFI contains, and who performs these actions. This data can be retrieved as historical records or in near real time, depending on the data flow configuration or pipeline schedule.

All extracted data required to construct the event logs is then stored in a relational SQL Server database. The underlying relational schema related to RFI management is illustrated in the ERD shown in Figure 6-2. Accordingly, enabling process awareness requires the transformation of this SQL data structure into an event log format. To achieve this, the derived transformer Figure 5-4 was applied to convert the SQL data records into process-aware event logs, in line with the approach in [207]. Figure 6-7 shows an example of the constructed event log for the RFI management process of a major commercial construction project, after applying this transformation.

RFI_ID	concept:name	org:role	time:timestamp
14855608	GC Creates Internal RFI Draft	8771117 - Assistant Project Manager - GC	2022-09-30 22:10:12
14855608	GC Creates RFI Draft in Procore	8771117 - Assistant Project Manager - GC	2022-09-30 22:10:22
14855608	GC Initiates & Submit/Forward to OR	8771117 - Assistant Project Manager - GC	2022-09-30 22:10:42
14855608	OR Reviews RFI and Replies	1597042 - Architect - OR050	2022-10-04 23:40:00
14855608	GC Closes RFI after OR Direct Review	8771117 - Assistant Project Manager - GC	2022-10-06 14:37:03
14884717	GC Creates Internal RFI Draft	8771117 - Assistant Project Manager - GC	2022-10-05 13:26:48
14884717	GC Creates RFI Draft in Procore	8771117 - Assistant Project Manager - GC	2022-10-05 13:26:58
14884717	GC Initiates & Submit/Forward to OR	8771117 - Assistant Project Manager - GC	2022-10-05 13:27:18
14884717	OR Forwards RFI OR	2959387 - RFI_Reviewer - OR157	2022-10-05 19:35:34
14884717	OR or GC Reviews RFI and 1ST Reply - LOOP 2	2959387 - RFI_Reviewer - OR157	2022-10-05 19:35:34
14884717	OR or GC Reviews RFI and 2ND Reply - LOOP 2	8998707 - RFI_Reviewer - OR157	2022-10-11 13:47:45
14884717	OR or GC Reviews RFI and 3RD Reply - LOOP 2	8364565 - MEP Engineer - OR112	2022-10-17 14:37:35
14884717	GC Closes RFI OR - LOOP 2	8771117 - Assistant Project Manager - GC	2022-10-17 15:56:06
14901887	GC Creates Internal RFI Draft	8771117 - Assistant Project Manager - GC	2022-10-06 16:59:58
14901887	GC Creates RFI Draft in Procore	8771117 - Assistant Project Manager - GC	2022-10-07 17:02:37
14901887	GC Initiates & Submit/Forward to OR	8771117 - Assistant Project Manager - GC	2022-10-07 17:02:57
14901887	OR Reviews RFI and 1ST Reply - LOOP 2	1597042 - Architect - OR050	2022-10-11 22:10:48
14901887	OR Reviews RFI and 2ND Reply - LOOP 2	8771117 - Assistant Project Manager - GC	2022-10-12 20:30:33
14901887	OR Reviews RFI and 3RD Reply - LOOP 2	1597042 - Architect - OR050	2022-10-12 22:53:01
14901887	GC Closes RFI OR - LOOP 2	8771117 - Assistant Project Manager - GC	2022-10-17 14:05:52
14904992	SC Creates RFI Draft	5293508 - Project Manager - SC091	2022-10-06 10:04:07
14904992	SC Submits RFI	5293508 - Project Manager - SC091	2022-10-06 10:04:27
14904992	GC Creates RFI Draft in Procore	8771117 - Assistant Project Manager - GC	2022-10-06 19:23:33
14904992	GC Initiates & Submit/Forward to OR	8771117 - Assistant Project Manager - GC	2022-10-06 19:23:53
14904992	OR Reviews RFI and Replies GC or SC - LOOP 1	8771117 - Assistant Project Manager - GC	2022-10-19 19:47:35
14904992	OR Reviews RFI - 2ND Reply LOOP 1	8359020 - RFI_Reviewer - OR011	2022-10-20 20:02:21
14904992	GC Closes RFI OR - LOOP 1	8771117 - Assistant Project Manager - GC	2022-10-21 13:31:30
14916121	GC Creates Internal RFI Draft	8771117 - Assistant Project Manager - GC	2022-10-07 18:52:57
14916121	GC Creates RFI Draft in Procore	8771117 - Assistant Project Manager - GC	2022-10-07 18:53:07
14916121	GC Initiates & Submit/Forward to OR	8771117 - Assistant Project Manager - GC	2022-10-07 18:53:27
14916121	OR Reviews RFI and Replies GC or SC - LOOP 1	8771117 - Assistant Project Manager - GC	2022-10-12 20:33:02
14916121	OR Reviews RFI - 2ND Reply LOOP 1	8359020 - RFI_Reviewer - OR011	2022-10-17 21:10:55
14916121	GC Closes RFI OR - LOOP 1	8771117 - Assistant Project Manager - GC	2022-10-21 13:24:24
RFI Instance ID	RFI Process Activity	Resource Role Actor	Time Stamp

Figure 6-7. Excerpt of an RFI management event log in CSV tabular format.

As a result of the comprehensive literature review presented in Chapter 3. , it was observed that, to date, no process mining–based studies have analyzed the actual performance of the RFI management process in real-world construction projects, particularly including the cost dimension from a process-oriented perspective. Therefore, a third algorithm, Figure 5-5, was developed in this study with a twofold purpose: (i) to add the cost dimension to the constructed RFI event logs as shown in Figure 6-8 ; and (ii) to serialize these process event logs into an XSD machine-readable structure to ensure their quality and conformance with the international XES event log standard [207]. The cost of each RFI event was calculated based on the base hourly pay rate of each project stakeholder role involved in performing the RFI tasks, multiplied by the service task time durations computed from the actual logged timestamps. It is important to note that process task durations were computed using the standard Québec working calendar for 2026. Adherence to the XES standard ensured consistent timestamp formatting based on ISO 8601 [91, 126], enabled the incorporation of domain-specific extensions, and supported the generation of a well-structured event log. The cost dimension was incorporated using XESame [114] and PM4Py [98], while RFI-sensitive content was anonymized to preserve stakeholder confidentiality.

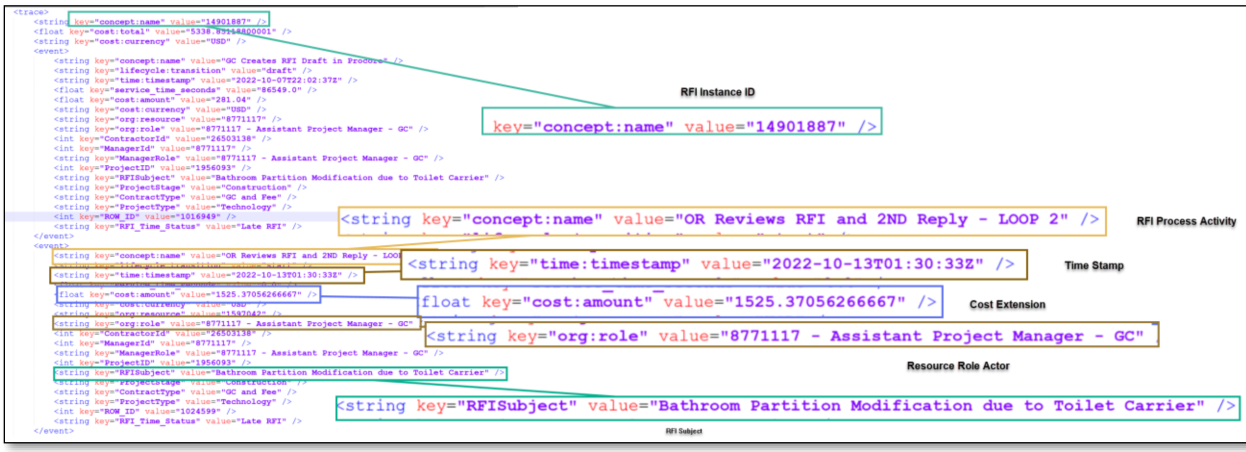


Figure 6-8. Excerpt of an RFI management event log in XES format enriched with cost dimension.

6.2.2. Process discovery and performance assessment – Level 4 (L4)

The systematic approach presented in Section 5.4.1 for the automated extraction of RFI operations data and event log generation was applied to the 71 selected construction projects and their 5,564 RFIs. Most construction organizations are project-driven by nature. While this is not necessarily a drawback, a major downside is the lack of visibility into the actual performance of core business processes that must be executed efficiently to deliver successful projects to clients, specifically on time and within budget. To address this limitation, a comprehensive quantitative performance assessment of the RFI management process was conducted across the 71 selected projects. From this analysis, a high-level insight into the overall health of the RFI process was extracted to support automated performance monitoring (i.e., PHM) within the GC's project portfolio. For this purpose, the weighted average process efficiencies in terms of time (WAPE) and cost (WACE) were calculated for each project using (Eq. 9) and (Eq. 10), respectively.

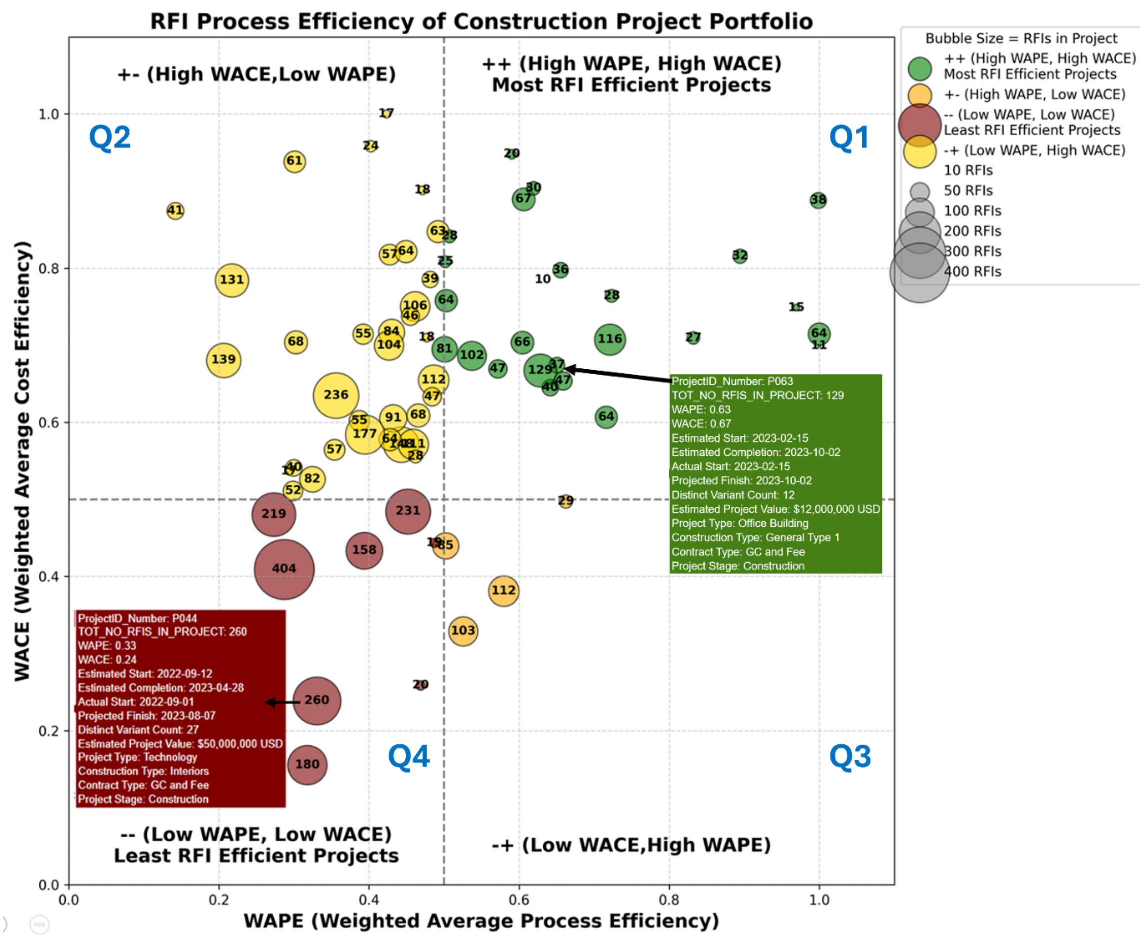


Figure 6-9. RFI process performance monitoring across project portfolio – process time-based efficiency (see interactive chart here)

Figure 6-9 provides an aggregated perspective of the RFI management process efficiency across the GC's project portfolio. Projects in the first quadrant (Q1) exhibit cost and time efficiency in managing RFIs. The second quadrant (Q2) represents projects with high-cost efficiency in RFI processing but low time efficiency. Projects in the third quadrant (Q3) showed relatively high time efficiency but low-cost efficiency. In contrast, the fourth quadrant (Q4) includes projects with poor efficiency in both cost and time. The time-based and cost-based process efficiency equations, (Eq. 9) and (Eq. 10), were introduced in Section 5.4.2.2. It is important to note that to compute the overall time-based process efficiency of each project, Lean-based metrics, i.e., cycle times, were computed for each of the actual RFI process instances executed within each project. Moreover, for the cost-based process efficiency computation, the cost threshold RFI was set to $2,500 USD based on the average cost of dominant (i.e., most frequent) process variants, which is within the ranges reported in the literature to date [227, 264]. However, this cost threshold can be defined by construction organizations according to their defined cost control strategies to be implemented to their core business processes.

To demonstrate the application of Modules M4 and M5 of the LPPM framework (see Figure 5-1) at the process level, and to identify performance issues in the execution of the RFI management process while exploring their potential root causes, a large commercial construction project was selected for further quantitative analysis. This project exhibited a high volume of RFIs and suboptimal performance, based on its computed WAPE = 0.33 and WACE = 0.24 process efficiency metrics within the project portfolio (see ProjectID = P044 in Figure 6-9). The selected project has an estimated value of $50 million USD under a cost-plus fixed fee contract arrangement.

6.2.2.1. Module 4 (M4): Process model discovery/extraction

One of the main applications of process mining is automated process model discovery. This task involves the extraction of actual process executions from event logs. In this study, the event log for the RFI management process of the selected commercial project (P044) was constructed in accordance with the XES standard [207] , as detailed in Section 5.4.1.3. The final event log consisted of 260 RFIs with a total of 1,660 recorded events. To enable the automated extraction of the RFI management process model for P044, two main process discovery techniques were adopted due to their reported benefits [126]. The Fuzzy Miner algorithm, described in Section 2.2.5, was applied to the constructed XES event log, as defined in (Eq. 4). This technique is widely used by process mining practitioners due to its simplicity and ease of interpretation. In addition to helping derive the RFI management process model, it provides a user-friendly representation as a DFG, which filters out less significant and weakly correlated events. Figure 6-10 shows the dominant E2E RFI management process model derived using Fuzzy Miner algorithm. The model was abstracted following Pareto principle, including only the top 20% of dominant process variants, which account for 80% of actual process executions. As highlighted in the callout, 67 RFIs were created by the SC, while 193 were issued directly by the GC and then officially assigned by the GC manager to an OR for review. Nine RFIs involved three rounds of back-and-forth communication, five others had four rounds, and 74 RFIs required escalations to multiple OR reviewers (nearly 30% of all issued RFIs).

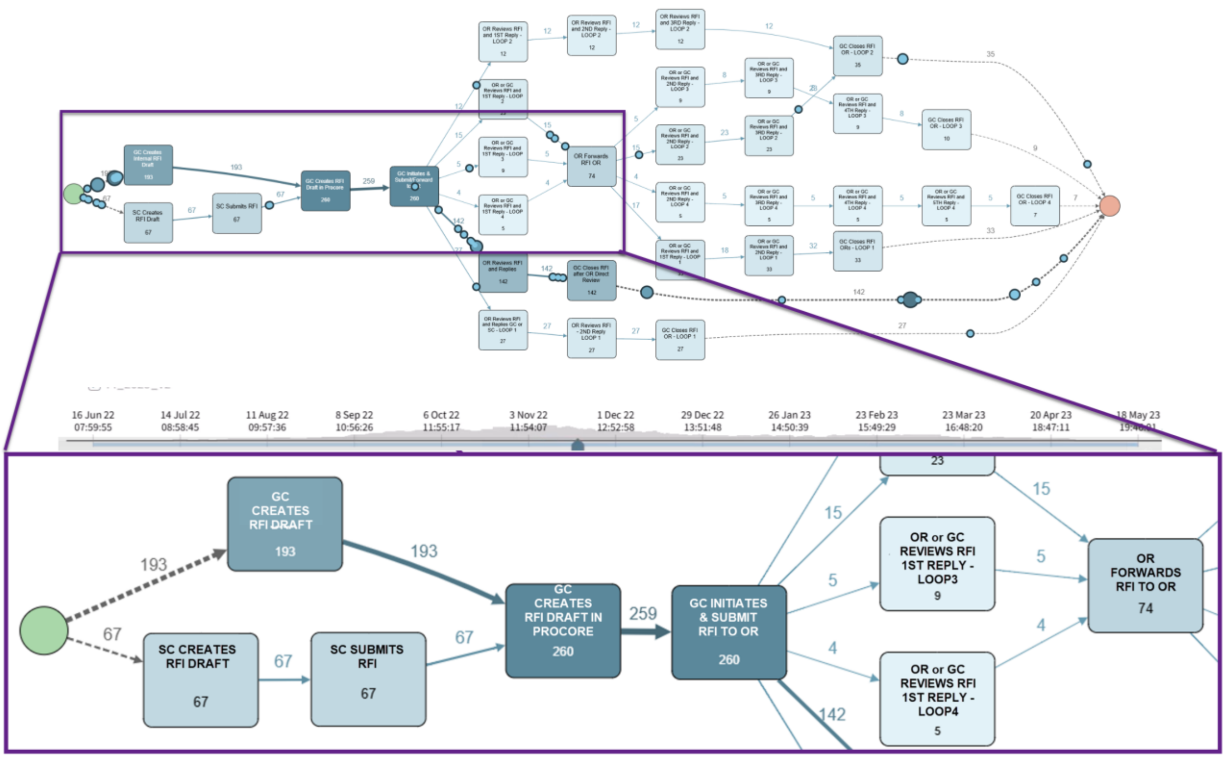


Figure 6-10. 'As-happened' RFI management process model derived with Fuzzy Miner.

Despite its benefits, the Fuzzy Miner algorithm is prone to produce unsound process models by either allowing extra unobserved behavior (divergence) or oversimplifying the process (convergence), potentially leading to misleading interpretations [91, 265]. Therefore, to ensure process model soundness as described in Section 2.2.5, the Inductive Miner algorithm was applied to the RFI management event log of the same project P044. A process tree with formal constructs was automatically derived as per (Eq. 5). As highlighted in the callout of Figure 6-11, RFI loops are represented in the derived process tree as a series of back-and-forth communication. That is, a single RFI can involve one or multiple reviews and follow-ups by different parties involved in the RFIs. In this case, the process is represented using sequential and exclusive choice relational constructs between activities, which also display their actual mean durations. For example, RFIs that require escalation to multiple OR reviewers (i.e., when an OR forwards the RFI to another OR) often involve one to three additional rounds of communication, thereby increasing their activity durations (i.e., darker blues) and the overall process lead time, as shown in Figure 6-11.

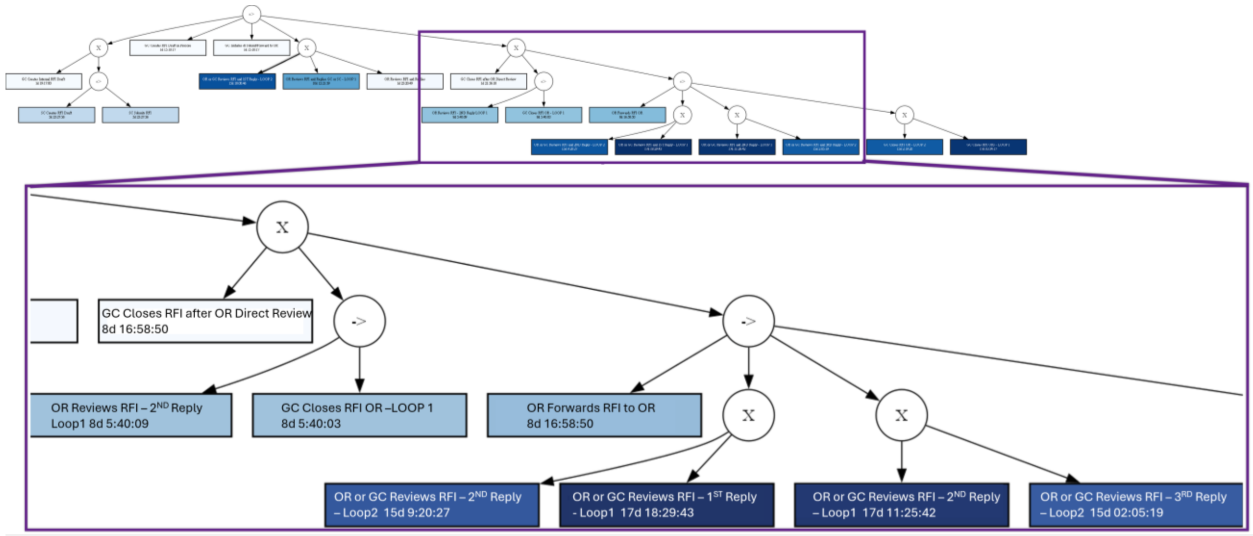


Figure 6-11. Automated extraction of RFI process model tree through Inductive Miner.

From a control-flow perspective, automated conformance checking was conducted using the Inductive Visual Miner and PM4Py to both visually and quantitatively assess the accuracy and quality of the derived dominant process model against the full RFI event log of project P044. The model demonstrated its best performance (i.e., its ability to represent reality) when 35% of infrequent process variants were filtered out. Model performance was assessed based on fitness and precision, as defined in (Eq. 7) and (Eq. 8).

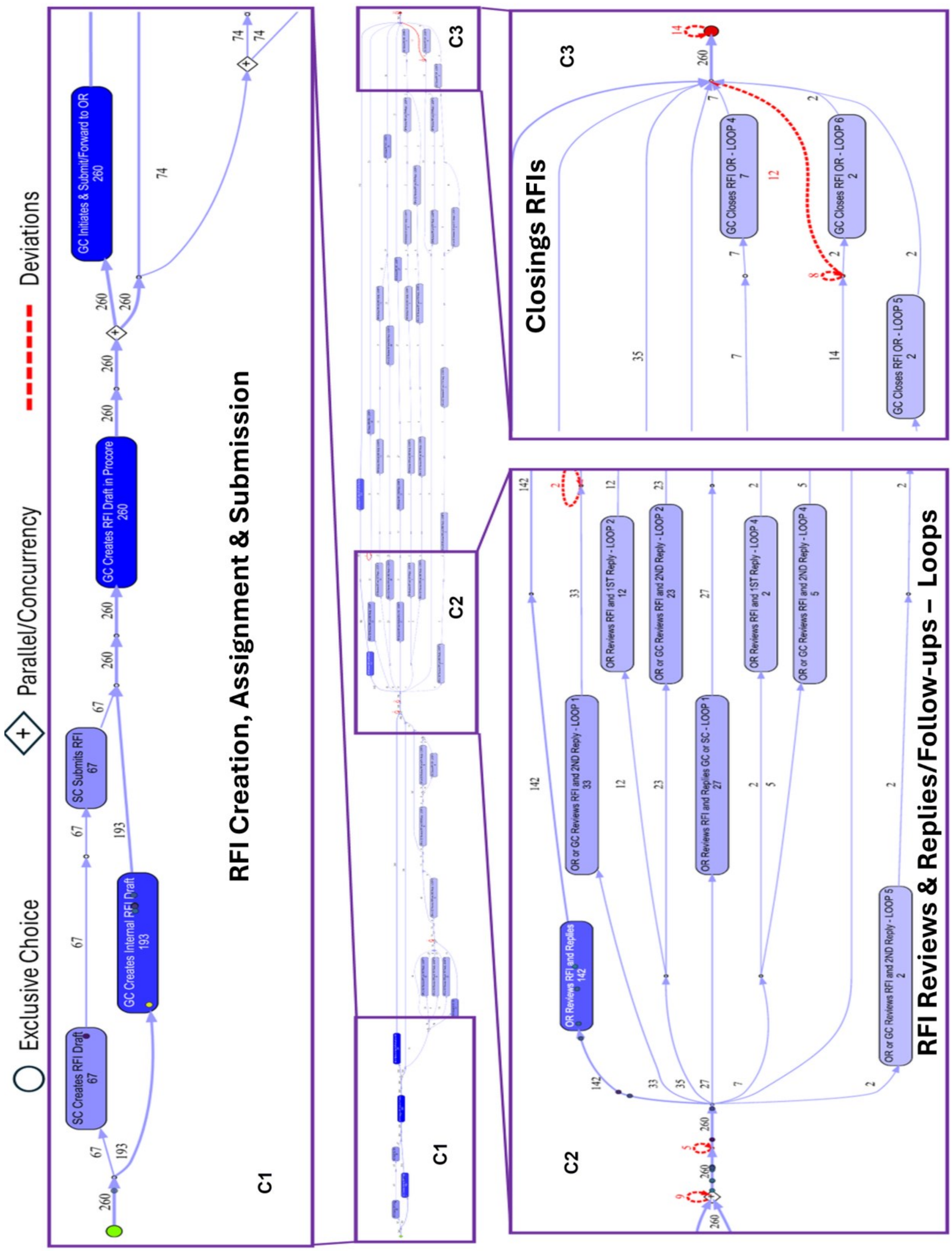


Figure 6-12. Automated process model conformance checking
Inductive visual Miner – Discovered dominant process model vs. RFI event log in P044

The assessment results are presented in Table 6-3 with a fitness value of 0.958, a precision of 0.761, and an F-score of 0.845, supporting the validity of the process model discovered. Callout 1 (C1) highlights the synchronous moves or full alignments, between the event log and the process model. The self-arcs in Callout 2 (C2) suggest either rework or additional loops present in the event log but not represented in the model. The circumventing arc in Callout 3 (C3) indicates deviations or skips where 12 out of 14 incoming RFIs were closed immediately after completing loop 3, while the remaining two required three additional rounds of reviews and replies as depicted in Figure 6-12.

Table 6-3. Process Model Discovery - Validation Results

Validation Metric	Value	Description
Synchronous moves	209	C1: Conformant activities and edges between process model and event log
Log Moves	39	C2: Rework or extra review loops (i.e., seen in the log but not in the model)
Model moves	12	C3: Skipped Activities (i.e., expected by the model but not seen in the log)
Fitness	0.958	≈ 1 model represents reality
Precision	0.761	≈ 0.8 model is neither too general nor too specific
F-score	0.845	Harmonic mean (i.e., > 0.8 robust valid model)

6.2.2.2. Module 5 (M5): Process Performance Assessment

Most business process improvement initiatives and process-oriented research in the construction domain, despite focusing on re-engineering, optimization, or automation, tend to overlook the essential prerequisites of discovering real-world E2E RFI process executions and quantitatively assessing their performance to identify and diagnose key inefficiencies. Without fulfilling these prerequisites, decision-making remains intuitive, judgment-based, and disconnected from factful insights. To address this aspect, the proposed PHM framework focuses on the automated quantitative performance assessment of the RFI process, leveraging the LPMM to uncover actual operational behavior and team dynamics, thereby enabling more targeted, evidence-based, and effective improvements. The RFI process assessment encompasses four core dimensions: time, cost, control-flow, and organizational perspectives.

Time perspective – First, time-oriented bottlenecks were identified from the automatically derived process model that captures the actual E2E RFI workflow executions for P044. These bottlenecks correspond to the activities and paths with the longest durations observed during RFI process execution for P044 as shown in Figure 6-13. For instance, the figure shows that the longest path between process activities occurs when RFIs are forwarded by the GC or by an OR to additional reviewers. This pattern occurred 74 times, with a batching subset of 21 RFIs taking an average of eight days solely for the forwarding step. Additionally, 13 RFIs experienced prolonged initial reply activity exceeding seven days by either the GC or the OR, which was followed by forwarding the RFIs to additional OR(s), resulting in multiple back-and-forth communications (i.e., loops 3, 5, and 8). Additionally, 41 RFIs experienced long waiting times of more than five days for the SC or GC to respond back to the OR with the necessary follow-ups.

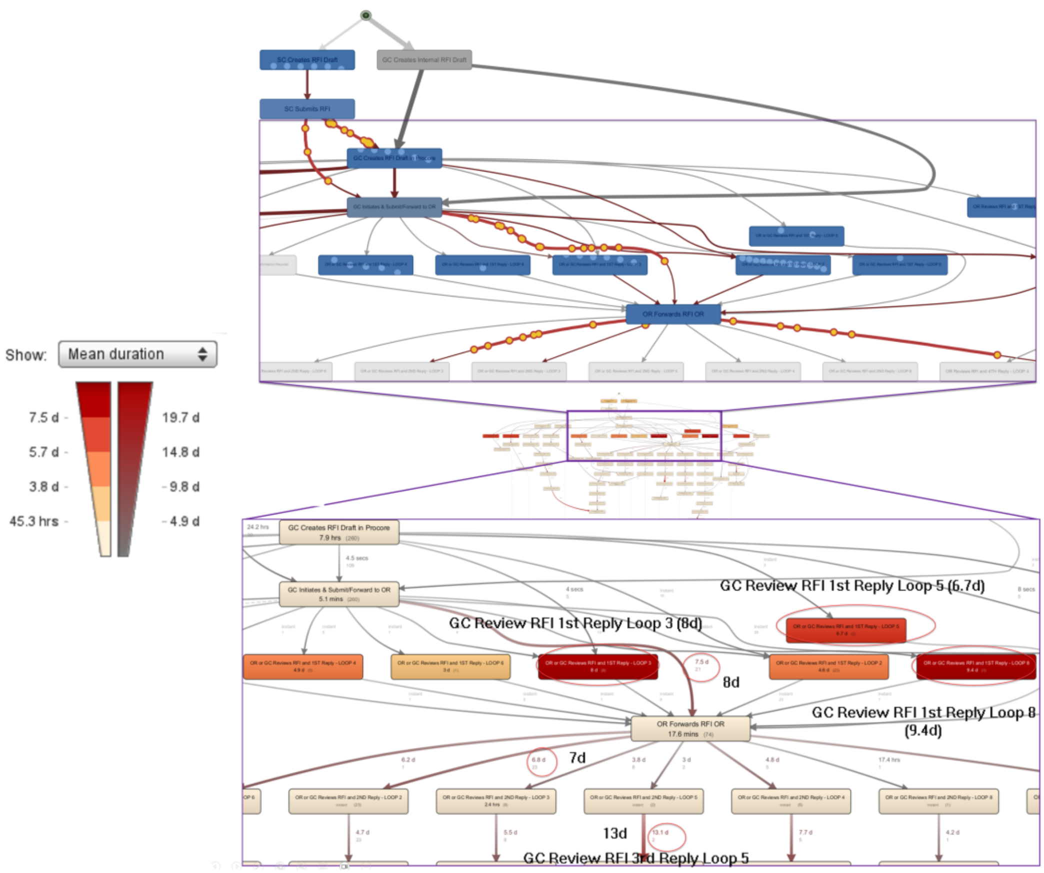


Figure 6-13. Automated identification of time-oriented bottlenecks in the RFI process.

To address the descriptive time-related questions outlined in Table 5-1, a cross-case analysis was conducted by comparing project P044, which exhibited suboptimal RFI performance, with project P063 (see Figure 6-9), which showed better RFI process efficiency. Both projects are commercial office buildings following a DBB delivery system, involving the same cost-plus fixed fee contract, and both are currently in the construction phase. However, they differ in terms of ownership and project size: P044 is developed by a technology-based client with an estimated project value of $50 Million USD., whereas P063 is managed by an insurance-based client with an estimated project value of $12 Million USD.

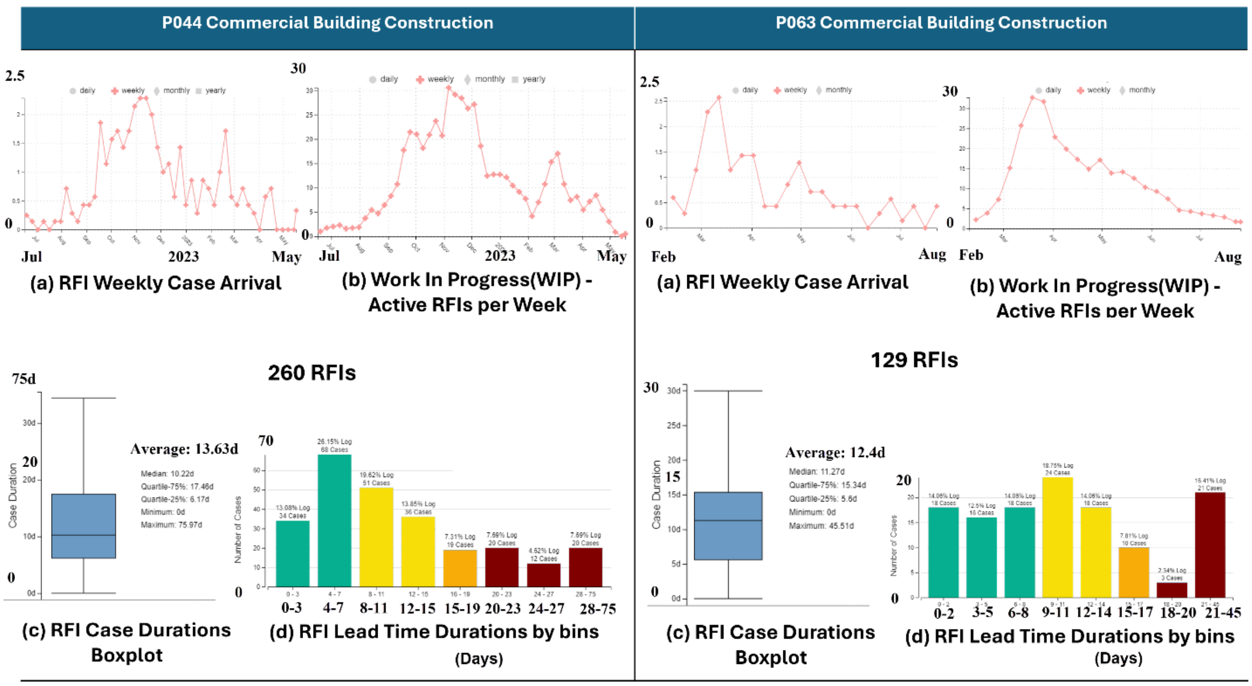


Figure 6-14. Cross-case analysis: time-oriented perspective of RFI process executions.

The following key findings can be observed in Figure 6-14 which presents the results of the overall comparative analysis.: (a) The RFI case arrival rate for both projects ranged between one and three RFIs per week; however, the line chart for P044 is skewed to the right, while that of P063 is skewed to the left, suggesting that RFI issues in P063 were raised more promptly and proactively from the outset of the construction phase; (b) The Work-In-Progress (WIP) analysis indicates that project stakeholders in P063 were able to address, resolve, and close RFIs more efficiently than those in P044; (c) The average resolution time per RFI was 13.63 days for P044 and 12.4 days for P063; (d) Although the total number of RFIs in P044 is double than that of P063, when adjusted for project size in terms of contract value, P044 received 5 RFIs per US$1 million, whereas P063 received 11 RFIs per US$1 million. Despite receiving more RFIs per unit cost, P063 resolved approximately 40.3% of RFIs within the 7-day due date, compared to 39.2% for P044 (see green bars in chart d). Last but not least, a greater number of RFIs in P044 experienced longer delays, i.e., late RFIs, than those in P063 (see yellow, orange, and red bars in chart d).

<u>*Cost perspective –*</u> To illustrate the process assessment from a cost perspective within the main case study project (P044), we automatically derived the dominant RFI process model at a 65 percent level of abstraction by excluding 35 percent of infrequent behavior. Similarly, construction organizations can derive their own accepted benchmark RFI process models and set cost thresholds aligned with their operational goals. This reference model is then compared against actual RFI process executions at an 85 percent level of abstraction, where frequency constraints are relaxed to identify cost overruns resulting from deviations from the ideal RFI process execution.

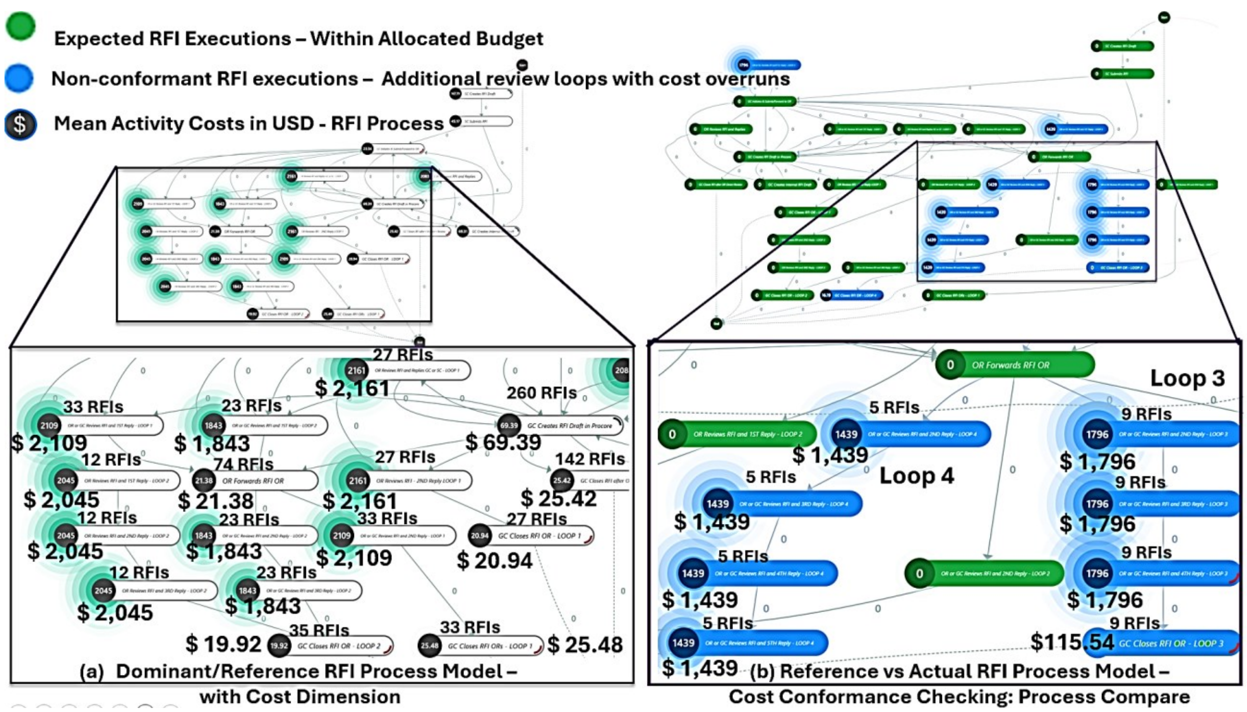


Figure 6-15. Automated Cost Conformance Checking of RFI Process in P044.

From Figure 6-15, it can be observed that: (a) the reference model presents each RFI process activity along with its concurred cost, reflecting the number of RFIs that pass through each activity; and (b) two additional loops (Loops 3 and 4) reveal undesirable extended RFI reviews, along with the corresponding delta average costs when compared to the benchmark model. The costs of the 'review' activity were assumed to be equally distributed based on the number of required RFI communication follow-ups and their computed durations across the process timeline.

For example, the cumulative cost for loop 3, involving four review follow-ups in a single RFI process trace execution, reached $7,300.00 USD. Similarly, in loop 4, with five review follow-ups, the total cumulative cost was $7,218.66 USD per RFI. In both cases, the cost is nearly three times higher than the maximum expected threshold of $2,500 USD per RFI. This suggests that, in loop 3, the same budget could have covered the review of approximately 27 RFIs instead of just nine, if the review process variation and duration had been better controlled. Likewise, in Loop 4, the same amount could have supported the review of 15 RFIs instead of just five or could have generated nearly $70,000 USD in savings by better controlling the RFI process variations associated with non-conformant Loops 3 and 4.

Control-flow perspective – Gaining visibility into the actual variability of core business process executions is essential for assessing, monitoring and controlling their performance. In this vein, a cross-case process variant analysis was carried out to compare the RFI process variations in project P044 with those in project P063. A process variant is a unique trace found in the constructed event logs (i.e., a distinct sequence of event activities). These variations often emerge due to factors such as the technical complexity of the RFI inquiry, the clarity of the questions posed, staffing availability, resource capacity, and the misrouting of RFIs due to an unclear definition of reviewer authorities, among other contributing aspects. Figure 6-16 shows the 'as-happened' RFI process execution variations for both commercial building projects, i.e., P044 and P063, which were automatically extracted from their corresponding event logs. As shown in Figure 6-16, P044 exhibited greater variability in the execution of the RFI management process than P063, with 28 variants compared to 19, respectively.

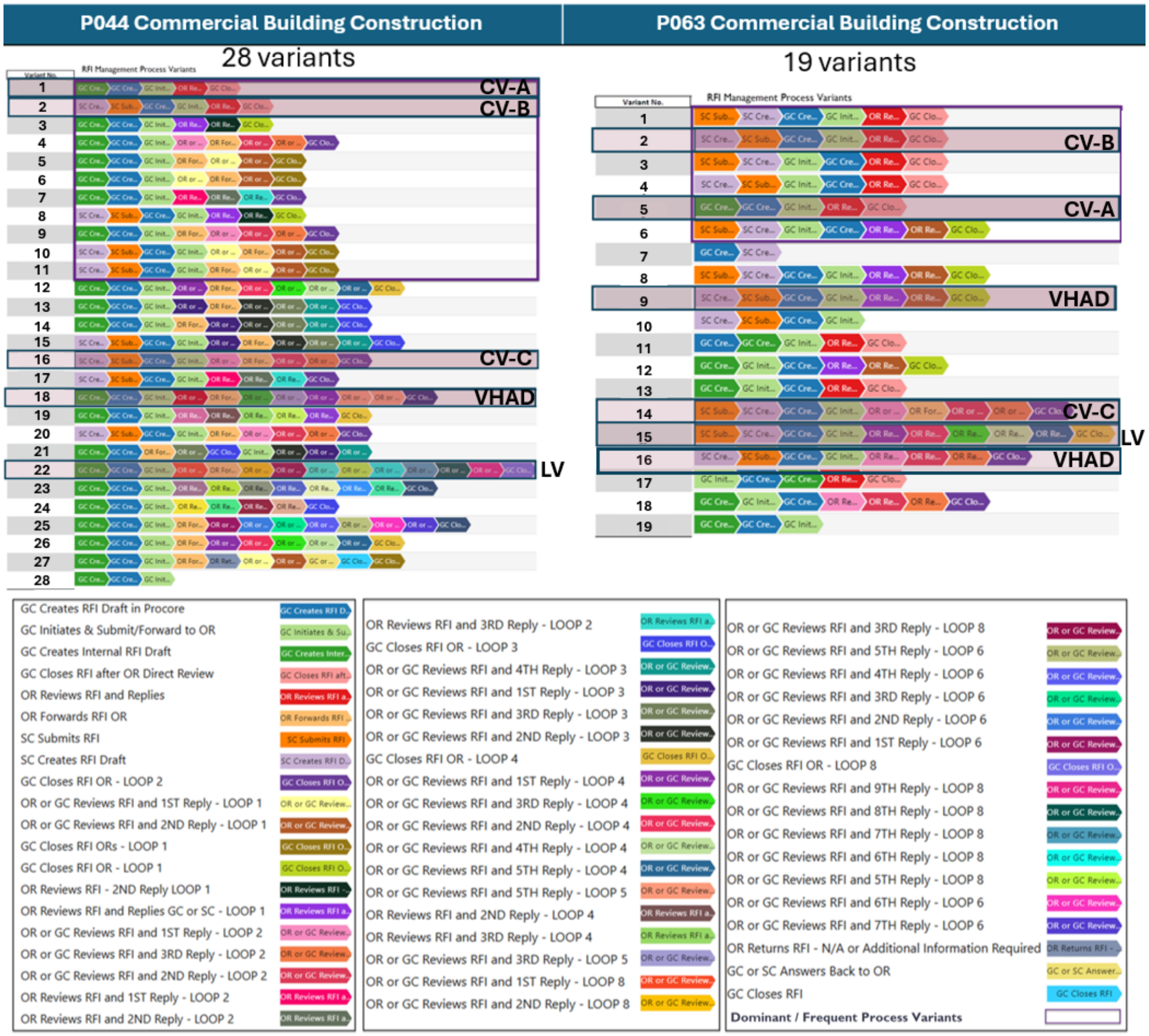


Figure 6-16. Cross-case variant analysis: 'As-Happened' RFI Executions. Common variants (CV-A, CV-B,CV-C); variants with high average durations (VHAD); longest variant (LV).

Ideally, the GC, who manages both projects, aims at resolving and closing any RFI within seven days once the RFI is received and initiated, while its maximum defined deadline is fourteen days. Considering this aspect, on the one hand, Figure 6-16 and Figure 6-17 shows the common variants whose RFIs followed the similar E2E execution sequences in both projects, namely CVA, CVB, and CVC. In addition, the most frequent variants for each project are also highlighted, specifically the first eleven for P044 and the first six for P063. Although, on average, both types of variants were resolved within the maximum threshold defined by the GC, several RFIs following these variants still experienced delays, with process lead times exceeding 14 days as shown in Figure 6-14.

P044 Commercial Building Construction

	Variant No.	Count	Event Count	Average Lead Time Duration	Finance Total $USD
CV-A	1	101 Cases 38.85 % Log	5	1w 1d 5h 13m 4s	228K
CV-B	2	41 Cases 15.77 % Log	6	1w 4d 20h 12m 58s	102K
	3	18 Cases 6.92 % Log	6	2w 10h 15m 23s	77K
	4	13 Cases 5.00 % Log	8	2w 2d 7h 17m 2s	76.1K
	5	12 Cases 4.62 % Log	7	1w 5d 16h 23m 45s	51.6K
	6	10 Cases 3.85 % Log	7	1w 4d 15h 29m 19s	41.7K
	7	10 Cases 3.85 % Log	7	2w 5d 15h 45m 58s	60K
	8	9 Cases 3.46 % Log	7	2w 3d 2h 1m 21s	45.2K
	9	7 Cases 2.69 % Log	8	3w 2d 17h 23m 58s	41.3K
	10	5 Cases 1.92 % Log	8	2w 3d 8h 44m 21s	19.8K
	11	5 Cases 1.92 % Log	8	1w 5d 4h 26m 18s	29.7K
	12	4 Cases 1.54 % Log	10	1mo 1w 5d 18h 37m 1s	32.2K
	13	3 Cases 1.15 % Log	9	1mo 11h 38m 55s	23.5K
	14	3 Cases 1.15 % Log	9	2w 6d 13h 44m 26s	25.8K
	15	2 Cases 0.77 % Log	10	3w 2d 7h 26m 25s	19.7K
CV-C	16	2 Cases 0.77 % Log	9	1w 3d 1h 33m 32s	16.6K
	17	2 Cases 0.77 % Log	8	1mo 2w 22m 37s	5.5K
VHAD	18	2 Cases 0.77 % Log	11	1mo 3w 2d 6h 5m 43s	16.7K
	19	2 Cases 0.77 % Log	9	2w 6d 8h 41m 20s	16.6K
	20	1 Case 0.38 % Log	9	3w 4d 7h 18m 58s	240.9
	21	1 Case 0.38 % Log	9	21h 45m 26s	11.4K
LV	22	1 Case 0.38 % Log	14	1mo 5d 17h 30m 18s	10.8K
	23	1 Case 0.38 % Log	11	1mo 2w 3d 15h 6m 24s	10.9K
	24	1 Case 0.38 % Log	8	3w 3h 14m 43s	4.5K
	25	1 Case 0.38 % Log	12	1mo 4d 21h 55m 15s	10.5K
	26	1 Case 0.38 % Log	10	1w 3d 3h 54m 59s	7.2K
	27	1 Case 0.38 % Log	10	2w 1d 22h 53m 21s	8.1K
	28	1 Case 0.38 % Log	3	15m	132.9

P063 Commercial Building Construction

Variant No.	Count	Event Count	Average Lead Time Duration	Finance Total $USD	
1	29 Cases 22.66 % Log	6	2w 1d 23h 46m 52s	112K	
2	28 Cases 21.88 % Log	6	1w 4d 17h 9m 31s	72.5K	CV-B
3	24 Cases 18.75 % Log	6	1w 1d 20h 12m 38s	58.5K	
4	12 Cases 9.38 % Log	6	1w 5d 9h 30m 22s	37.9K	
5	8 Cases 6.25 % Log	5	2w 1d 18h 31m 18s	26.8K	CV-A
6	4 Cases 3.13 % Log	7	2w 21h 24m 19s	14.9K	
7	4 Cases 3.13 % Log	2	1d 12h	426.9	
8	3 Cases 2.34 % Log	7	2w 5d 3h 29m 12s	12.5K	
9	2 Cases 1.56 % Log	7	3w 2d 22h 53m 1s	16.5K	VHAD
10	2 Cases 1.56 % Log	4	1d 40m 10s	1.5K	
11	2 Cases 1.56 % Log	5	2w 4d 25m 29s	7.2K	
12	2 Cases 1.56 % Log	6	1w 1d 22h 23m 22s	7.8K	
13	2 Cases 1.56 % Log	5	1w 5d 14m 28s	970.7	
14	1 Case 0.78 % Log	9	2w 3h 34m 28s	6.1K	CV-C
15	1 Case 0.78 % Log	10	2w 1d 6h 40m 56s	5.5K	LV
16	1 Case 0.78 % Log	8	3w 2d 22h 37m 16s	8.1K	VHAD
17	1 Case 0.78 % Log	5	1d 19h 32m 47s	567.84	
18	1 Case 0.78 % Log	7	17h 46m 27s	440.16	
19	1 Case 0.78 % Log	3	10m	76.23	

Figure 6-17. Cross-case variant performance analysis : As-Happened' RFI Executions. Common Variants (CV-A,CV-B,CV-C); Variants with high average durations (VHAD); Longest variant (LV).

On the other hand, the longest variants (LV), as well as those with high average durations, were also identified in both figures. For instance, in P044, variant 22 required 14 process activities, whereas in P063, variant 15 involved 10 steps. Both with multiple loops of back-and-forth communications. In addition, process variant 18 in P044 and variants 9 and 16 in P063 exhibited high average durations. This second type of variation, although less frequent, resulted in significant time and cost overruns. From a quantitative standpoint, considering maximum thresholds of 14 days for time and $2,500 USD for target cost, the infrequent variants, excluding those whose RFIs were never answered (i.e., variant 28 in P044 and variants 7 and 19 in P063), account for a cumulative time overrun of approximately 7.4 months of delayed RFIs for P044 and one month for P063.

Moreover, potential savings of approximately $87.9 K USD for P044 and $22.2 K USD for P063 could have been realized if the RFI process executions within these infrequent variants had been better controlled to remain within the established thresholds. This could be achieved not only by reducing the number of process variations, but also by implementing targeted process improvements, such as clearly defining reviewer authorities to avoid unnecessary activities like multiple forwards to multiple reviewers (e.g., the "OR forwards RFI to OR" activity shown in Figure 6-16). These undesired activities are among the bottlenecks identified in Figure 6-13, which have substantial adverse impacts on time, cost, and efficiency, potentially incurring additional sanction fees for delayed RFIs, if such penalties are contractually stipulated between the involved stakeholders.

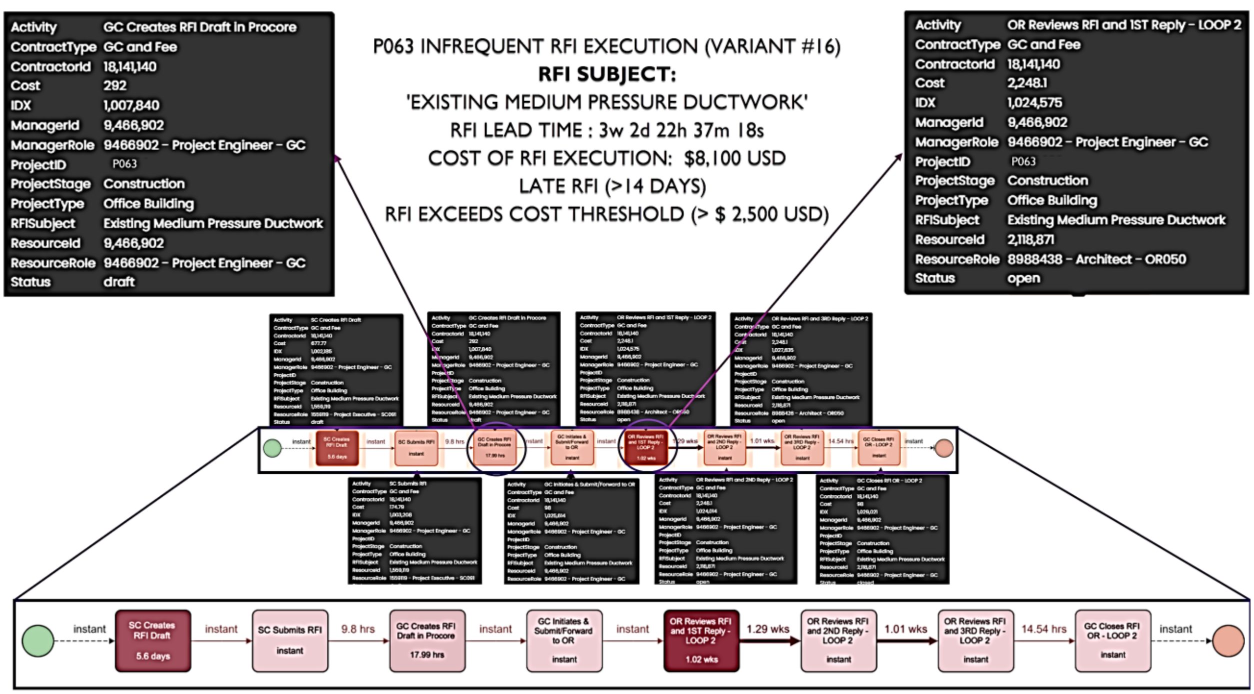


Figure 6-18. Non-conformant 'As-happened' E2E RFI process model execution of variant #16 in P063.

As shown in Figure 6-18, variant #16 from P063 is a non-conformant RFI process execution that exceeds the GC's maximum time and cost thresholds of 14 days and $2,500 USD, respectively. This execution suffers from various bottlenecks in the review process with multiple back-and-forth communication loops. More specifically, the second loop in variant #16 means that two additional review follow-ups occurred after the first reviewer's response, only then, when neither the GC nor the OR have any further requirements and the SC has accepted the given response, does the GC proceed to close the RFI. These aspects result in significant time overruns and in additional costs being incurred, as depicted in Figure 6-18.

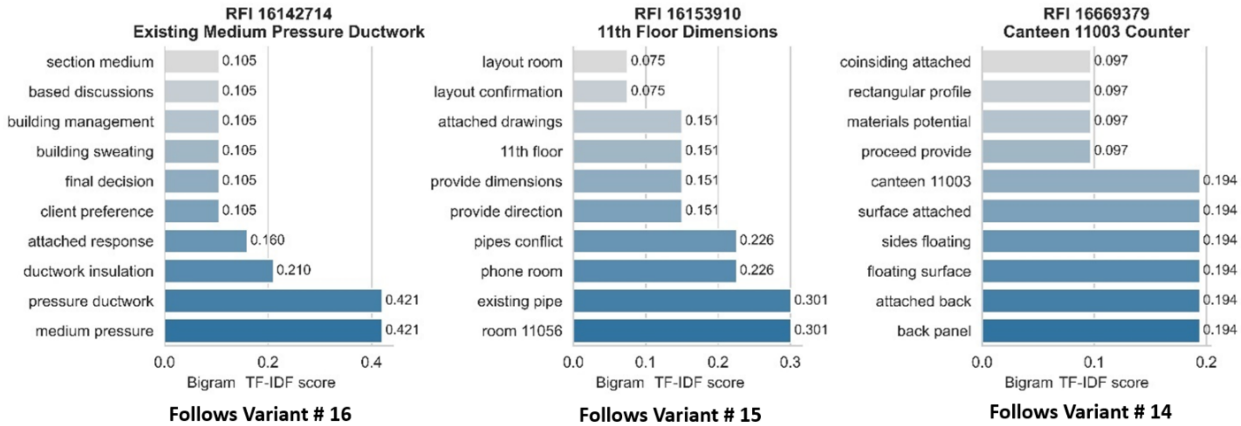


Figure 6-19. Bigram analysis of RFIs following non-conformant process variants in P063.

A Term Frequency–Inverse Document Frequency (TF–IDF) analysis, a text mining technique, was applied to identify the most frequent and distinctive bigrams within the content of each of the three RFIs associated with impactful variant executions in P063. To provide broader context, the actual execution of RFI (ID 16142714), which followed variant #16 (see Figure 6-18), was originated by the mechanical SC who observed on site that several sections of air ventilation ducts across multiple floors were not insulated, as the design drawings indicated no insulation for those sections. On 6 March 2023, the SC field engineer submitted a formal RFI to the consultant OR, requesting confirmation on whether these sections should remain uninsulated or be added to the scope. Eight days later, the first response from OR1 (architectural consultant) confirmed that the drawings showed no insulation but referred the RFI issue to both the facility manager (FM) and the mechanical engineering consultant. Nine days after this, the OR2 (FM) advised that those specific sections were unlikely to sweat and could remain uninsulated but deferred the "final decision" to the mechanical consultant (OR3), citing it as a "client preference." Sixteen days after the first response, OR3 instructed the SC to insulate the identified sections on all floors to comply with city code. In total, the decision took more than 24 days, disrupting and delaying on-site work as well as incurring additional RFI management costs. Similarly, the textual context for each of the three less frequent yet impactful RFI variants is reflected in the identified bigrams shown in Figure 6-19.

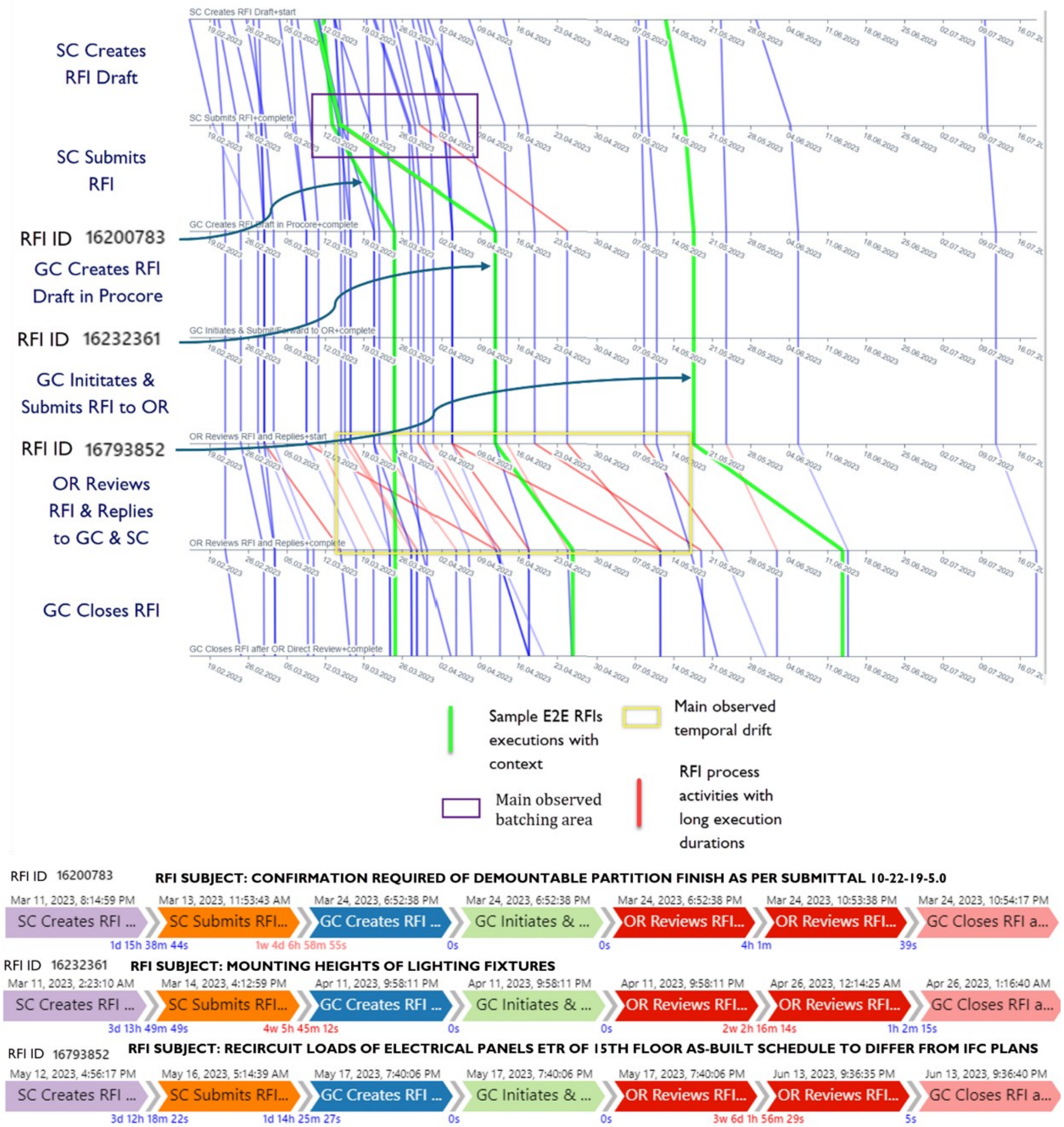


Figure 6-20. Performance Spectrum Analysis: Batching and Drift Patterns Detection P063.

<u>*Control-flow perspective –*</u> Figure 6-20 shows the application of the process performance spectrum technique [224, 238] to automatically detect main process drifts and batching points over time. In this approach, each RFI following execution variant #2 within P063 is plotted in the chart. Each line represents a single RFI, traced from creation at the top, through all intermediate activities, to closure at the bottom, while the horizontal (x) axis indicates the period during which the RFI was processed. In this case, the RFI submission from the SC, the RFI creation by the GC, and the RFI reviews by the OR are the main activities that tend to result in longer process delays.

RFIs following variant #2 are always initiated by the SC, often derived from a field observation or correspondence in Procore or by email communication. The RFI draft creation involves identifying the issue and drafting the request, after which it is submitted to the GC's RFI manager for validation and official creation in Procore. The first major batching point appears between March 12 and April 2, 2023, when most RFIs, arriving in the project's early phase (see Figure 6-14), were queued until the GC created the official drafts in Procore. This batching period also affected review durations, contributing to a clear temporary drift in RFI reviews between March 12 and May 14, 2023. For instance, the RFI 16200783 whose content was simply requesting the OR confirmation about the wall partition finish was less impacted by this batching when compared to the RFI 16232361 which was requesting confirmation about the high for lighting fixtures as it was not specified on the Issued for Construction drawings.

However, batching was not the only factor influencing review durations; other aspects also played a significant role. For example, RFI 16793852, which was outside the batching period, was affected by two main factors. First, it was technically complex, as on-site conditions on the 15th floor and the number and use of various equipment required a rearrangement of the electrical loads supplied by the panels. Second, this RFI, submitted by the SC to the GC's specialist on May 17, did not receive a reply, possibly due to staff unavailability or unclear reviewer and approval authorities. Due to its urgency, on May 24, the SC submitted a second RFI with the same question. This second RFI was forwarded to the electrical engineering consultant, who provided a response 18 days later, detailing the new electrical load distribution. Based on this response, the GC issued an official reply for both RFIs, instructing the SC to follow the OR's directions and subsequently closed them.

Organizational perspective – Construction projects, by nature, demand the continuous involvement of various multidisciplinary teams, whose close collaboration requires constant, effective communication and frequent information exchanges among stakeholders through the entire project lifecycle to ensure successful execution and delivery. In this regard, a well-defined RFI management process, with clear identification of the different parties involved in its execution, is of paramount importance for enabling and guiding this collaboration in a structured and efficient manner.

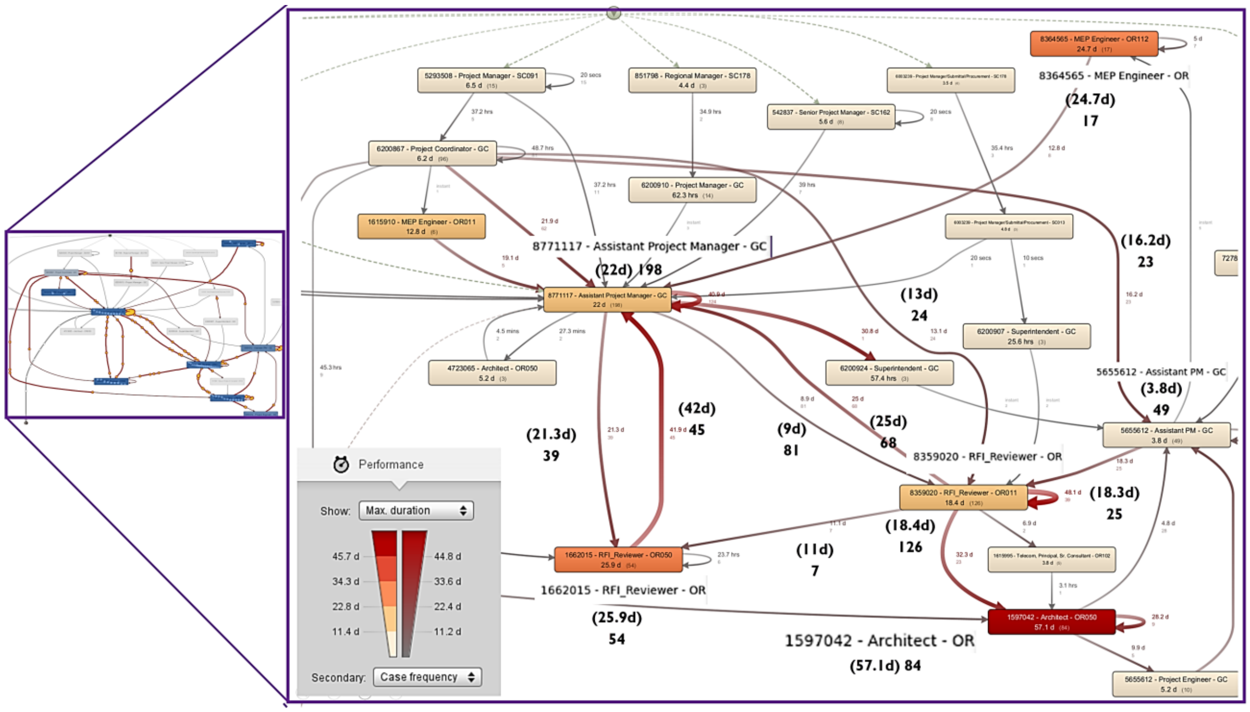


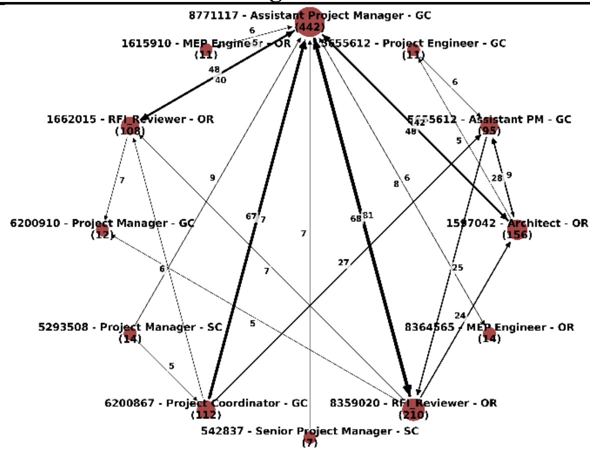
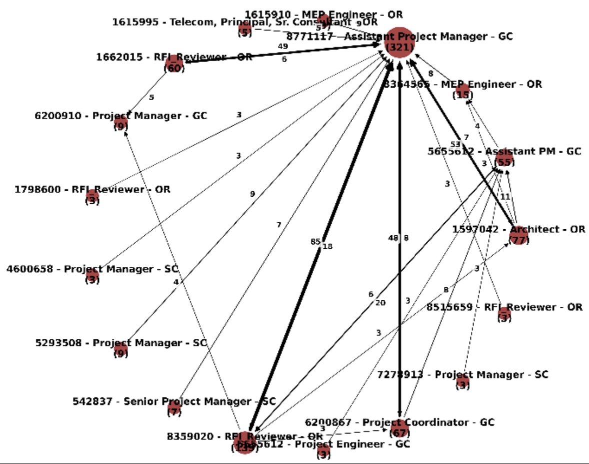
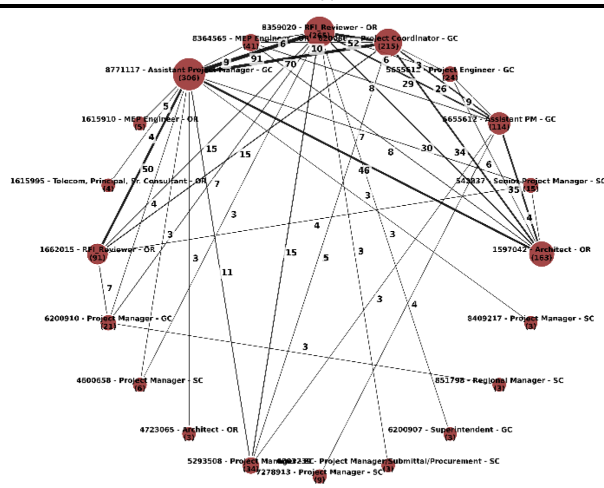
Figure 6-21. Work exchanges by stakeholder role in the RFI process, highlighting duration and frequency

In this context, the event logs described in Section 5.4.1.3 were enriched with the stakeholders' roles, aligned with their multidisciplinary functions in the GC's projects to analyze the inter-organizational team dynamics within the actual RFI process executions. Figure 6-21 shows the mining results considering this role perspective, representing the work, communication, and information exchanges among the stakeholders related to the execution of the RFI process in P044. The figure highlights performance in terms of the maximum duration of RFIs handled by specific project roles, as well as the frequency count of RFIs exchanged among these roles. It can be noted that the GC's Assistant Project Manager (ID = 8771117) carries the heaviest workload, handling approximately 76% of the RFIs (198) routed through this role, whereas another Assistant PM and the Project Coordinator are involved in only 96 and 49 cases, respectively. Furthermore, specific OR reviewers such as the Architect, the Mechanical Engineer, and the OR Consultants account for some of the longest durations in reviewing RFIs and providing responses to the GC. This does not necessarily imply inefficiency on the part of the role itself, but it does raise a warning about ineffective communication in the resolution process, such as multiple forwarding and back-and-forth exchanges. These factors affect the overall efficiency of the RFI process.

To further evaluate these team dynamics quantitatively, a social network analysis (SNA) was conducted based on the constructed RFI process event log for project P044. Three main social networks were derived based on [231] to identify the topological patterns of stakeholder interactions during the execution of the RFI process. These sociograms are represented as interconnected networks of actors, where each node corresponds to a stakeholder role participating in the RFI process, and the edges connecting the nodes represent different types of interactions or relationship patterns. The handover-of-work network was constructed by identifying directly-follows relationships, where one role's activity is immediately followed by another's, indicating causal dependency and transfer of work. The subcontracting network was derived by detecting dependency-chain patterns in which certain roles perform 'in-between' RFI activities that enable another actor to continue the process. The working-together network was generated by identifying co-occurrence patterns where two roles collaborate in the same RFI, regardless of temporal sequence or order dependency.

Table 6-4 presents the performance results of the three derived social networks as well as their main SNA metrics derived and calculated using PM4PY based on the approaches reported in previous studies [36, 90, 229]. First, the handover-of-work network and its computed metrics revealed 596 handovers among 16 roles (nodes) via 44 directed ties. The average weighted degree ≈ 82 handovers per role and the average unweighted degree ≈ 6 role-to-role interactions indicate high traffic with frequent back-and-forth ("ping-pong") exchanges. Consider, for instance, the 24 handovers between the OR Reviewer (ID 8359020) and the Architect. Moreover, density ≈ 0.183 reflects selective connectivity, neither highly dense nor siloed. The diameter is 4, so the longest shortest E2E RFI route requires four handovers. Four role clusters are observed (GC, SCs, specialty consultants, and ORs); however, modularity is low (Q ≈ 0.01), indicating both sparse community structure and multidisciplinary collaboration, which entails higher coordination overhead. In general, node size indicates each role's handover degree, edge width indicates the handover weight (traffic), and arrows show the direction of process flow.

Table 6-4. Quantitative SNA metrics of team dynamics in the RFI management process (P044)

Network	Discovered Organizational Network	Performance Metrics Results
(a) Handover of Work (see interactive chart here)		Number of Nodes: 16 Number of Edges: 44 Total Links: 596 Average Degree: 82.125 Network Density: 0.1833 Network Diameter: 4 Average Path Length: 1.8181 Number of Clusters: 4 Modularity: 0.01
(b) Subcontracting Network (see interactive chart here)		Number of Nodes: 17 Number of Edges: 28 Total Links: 394 Average Degree: 46.3529 Network Density: 0.1029 Network Diameter: 3 Average Path Length: 1.6904 Number of Clusters: 4 Modularity: 0.032
(c) Working Together (see interactive chart here)		Number of Nodes: 20 Number of Edges: 42 Total Links: 664 Average Degree: 66.4 Network Density: 0.2210 Network Diameter: 3 Average Path Length: 1.9789 Number of Clusters: 3 Modularity: 0.0711

Second, the subcontracting network comprises 17 roles linked by 28 directed ties, with 394 detected subcontracting patterns in which an intermediary role mediates RFI activity between one other role or two other roles. Relative to the handover network, it is sparser (density ≈ 0.1029 vs 0.1833), involves fewer role-to-role partners (avg unweighted degree ≈ 3.29 vs ≈ 6), and carries less traffic per role (avg weighted degree ≈ 46.35 vs ≈ 82). Paths are slightly shorter (average path length ≈ 1.6904; diameter 3 vs 1.8181 and 4), consistent with more streamlined, hierarchical exchanges over a limited set of channels. Four clusters are again present, and modularity remains low (Q ≈ 0.032, slightly higher than 0.01), indicating only modest structural separation with continued cross-role collaboration. As observed in the network, the Assistant PM on the GC side and the Architect OR consistently serve as in-between subcontracting intermediaries in the process.

Third, the working-together network spans 20 roles and 42 ties with 664 co-presence links, showing the broadest participation and highest connectivity (density ≈ 0.221) compared with the other two networks. Collaboration traffic per role is significant (average weighted degree ≈ 66.4. Distances remain short (average path length ≈ 1.9789; diameter 3). Fewer clusters are present (3), and modularity is the highest of the three yet still low (Q ≈ 0.0711), indicating limited intra-group cohesion with sustained inter-group interaction. Multiple collaborations are observed between the PC and PM (90/70) and between the PM and the Architect OR (46).

Each project role carries a distinct level of importance in RFI management; quantifying it is essential for identifying central, high-workload participants and planning resource allocation to streamline RFI execution. In this context, Figure 6-22 compares normalized PageRank, Hub, and Authority scores across the three RFI network layers, which were determined based on [90]. PageRank reflects how often a role is encountered when information moves through the network; high values indicate central, frequently visited roles. The OR RFI Reviewer (ID 8359020) shows the highest PageRank, especially in the working-together layer (≈0.85) and strongly in handover (≈0.60), indicating that most RFI episodes pass through this role. Hub (HITS hub) identifies dispatchers that point to authoritative deciders. The OR RFI Reviewer (8359020) and the GC Project Coordinator (ID 6200867) reach the top hub scores in handover and subcontracting (≈1.00), consistent with triage and routing functions. Authority (HITS authority) captures destinations that receive links from good hubs. The GC Assistant Project Manager (ID 8771117) holds the top authority scores (≈1.00 across layers), and the Architect (ID 1597042) is also prominent in working together (≈0.60), indicating frequent reliance on these roles for answers or approvals. Overall, the pattern reveals a hub-to-authority backbone: RFI reviewers and coordinators route queries, while the GC assistant PM and Architect are common endpoints, which explain short paths but also highlight potential bottlenecks at these roles.

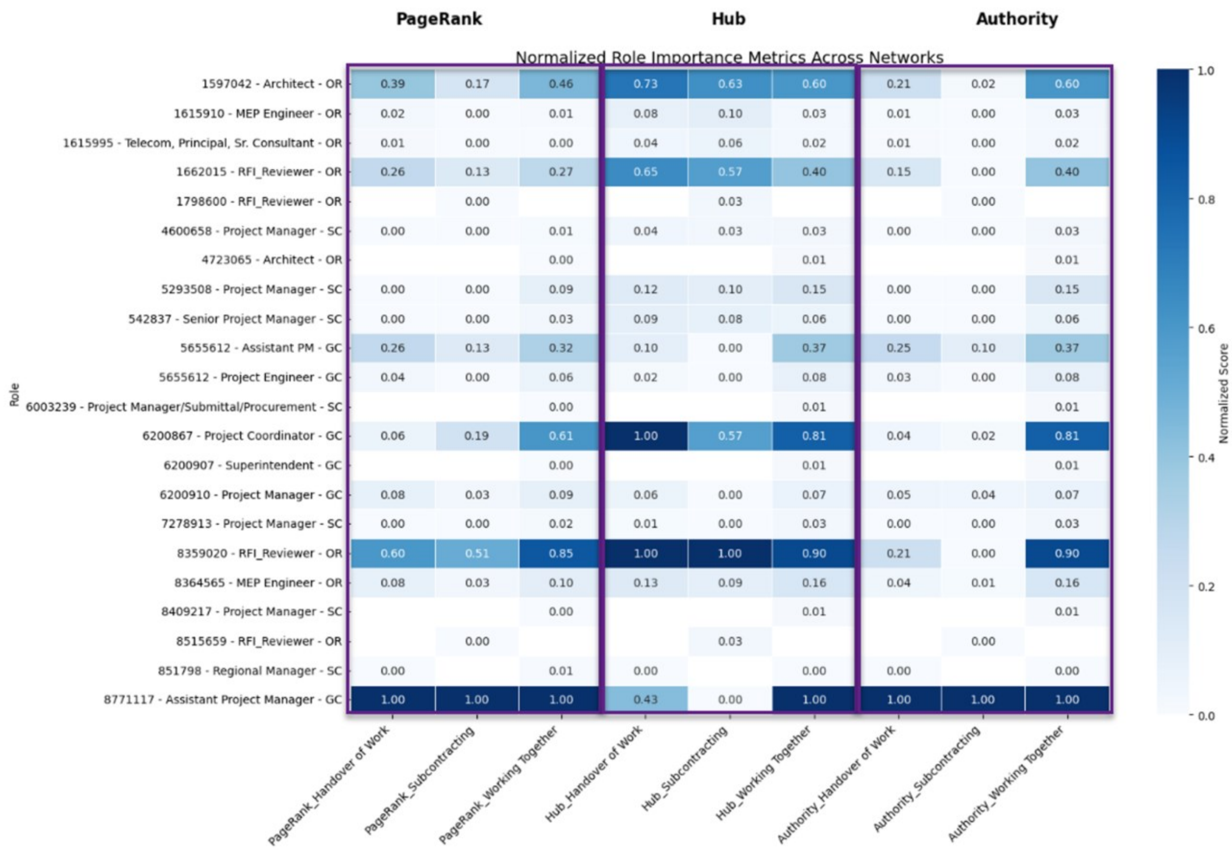


Figure 6-22. RFI process participants ranked by role importance across each social network.

6.2.3. Process-driven decision-making – Level 5 (L5)

6.2.3.1. Module 6 (M6): Process Improvement Opportunities

Actionable insights, in the form of prioritized recommendations for process improvement, were derived from the diagnostic results of the "as-happened" RFI process assessment conducted across the four-dimensional perspectives outlined in the previous sections. **Table 6-5** presents the summary of the data-driven results from the RFI process analysis. This summary provides visibility and transparency into the qualitative and quantitative performance of the actual E2E execution of the RFI process, outlining the activities performed, their sequence and duration, key process variations and bottlenecks, the roles and resources involved, workload distribution, and the root causes of delays, cost overruns, and efficiency losses.

Based on these data-driven results, 42 targeted process improvement recommendations were derived, with at least 10 suggested improvements identified for each dimensional perspective summarized in Table 6-6. For instance, under the control-flow perspective, the actual E2E RFI workflow exhibited high process variation, which underscores the need for automated RFI routing strategies, such as classification by type and priority, to maintain process stability and ensure a standardized flow that can be consistently reused across projects.

The derived process improvement recommendations were then prioritized from a construction business perspective using an impact–effort matrix as a decision-support tool broadly employed in process management to guide which initiatives should be implemented first to maximize value. The prioritization was informed by two dimensions: the relative level of implementation effort required for each action and the potential business value realized in terms of time and cost savings. This systematic evaluation not only facilitated the classification of recommendations into low-effort, high-impact interventions, medium-term actions, and long-term improvements, but also framed the prioritization primarily from a GC standpoint while remaining applicable to SC and OR parties involved in the RFI process.

Table 6-5. Data-driven aggregated results of "as-happened" RFI performance assessment via D2P method

Perspective	Question ID (Table 2)	Summarized Derived Insights from Data-driven Process Analysis	Perspective	Question ID (Table 2)	Summarized Derived Insights from Data-driven Process Analysis
Control-flow (CF)	Q1	The discovered 'as-happened' dominant RFI process model consists of 32 out of 75 distinct activities including four back-and-forth loops.	Cost (C)	Q1	For P044, 5.2 RFIs / $ MILLION USD For P063, 10.75 RFIs / $ MILLION USD
	Q2	For P044, WAPE = 0.33 For P063, WAPE = 0.63		Q2	For P044, $ 70,218 USD For P063, $ 53, 766 USD
	Q3	For P044, the first 11 variants out of 28 For P063, the first 6 variants out of 19		Q3	For P044, $ 992,774 USD For P063, $ 390, 282 USD
	Q4	The detected deviations included skipped RFI review activities or unanswered RFIs, while the main bottlenecks comprised lengthy reviews with various back-and-forth communication (i.e., more than 4 loops).		Q4	For P044, Variant 21 where a single RFI review incurred a cost of 11.4K For P063, Variant 16 where a single RFI review incurred a cost of 8.1K
	Q5	For P044, variants 18, 22, and 23 For P063, variants 9, 15, and 16		Q5	For P044, WACE = 0.24 For P063, WACE = 0.67
	Q6	In general, it was observed that process deviations, rework, and bottlenecks happened mainly due to the following reasons: Frequent value-engineering decisions with multiple alternatives; Multiple forwardings; Multiple decision-makers/approver instances; Batching RFIs; Ping-pong pattern (multiple back-and-forths).		Q6	Extra review loops 3 and 4 within P044 generated cost overruns of about $114,000 USD
Organizational (O)	Q1	Assistant PM GC (ID = 8771117); OR Reviewer (ID= 8359020) & PC GC (ID = 6200867)		Q7	Multiple forwardings; undefined reviewing authorities; poor accountability ping-pong patterns; multiple project roles involved the review process some without decision-making authority.
	Q2	Architect OR (ID = 1590742) & Assistant PM GC (ID = 8771117)		Q8	For P044, the first 11 variants out of 28 For P063, the first 6 variants out of 19
	Q3	Accountability ping-pong pattern detected multiple forwardings		Q9	1. RFIs submitted as additional scope detected or with cost impact specified. 2.Multiple redundant RFI submissions due to unanswered RFIs. 3. Multiple questions included within the same RFI. 4.Unclear questions or not applicable RFIs. 5. Undefined reviewing authorities. 6. RFIs with unreadable or unclear attachments. 7.. Looking for right attachment/information. 8 .Non-standardized naming conventions for attachments and documents. 9. Unknown requirements and/or specifications. 10. Inexperienced GCs, SC or ORs. 11. Poor design specifications. 12. Unavailable materials or components.
	Q4	Assistant PM GC (ID = 8771117) is central/critical resource in the RFI workflow experiencing higher work loads than for instance Assistant PM GC (ID= 5855612), which at the same time yields longer RFI resolution lead times.			
	Q5	Longer reviews are expected whenever the following roles are involved: The Architect OR (ID = 1590742), OR Reviewer (ID= 8359020); and the MEP Engineer OR (ID =8364585).			

Table 6-6. Actionable data-driven recommendations for RFI process improvement based on diagnostics.

Perspective	Targeted Data-driven Recommendations for RFI Process Performance Improvement	Perspective	Targeted Data-driven Recommendations for RFI Process Performance Improvement
Control-flow (CF)	**R1** Set thresholds for communication loops (<4) & number of involved reviewing authorities (2-3)	**Cost (C)**	**R1** Implement tiered-reviews by RFI type, RFI cost and roles
	R2 Implement an atomic RFIs strategy (i.e. single RFI question)		**R2** Escalate cost-sensitive RFIs to roles with sufficient decision-making authority.
	R3 Predefine reviewing authorities according to the RFI type received		**R3** Establish cost-sharing agreements for RFI management overheads arising from design errors or omissions
	R4 Set RFI prioritizations by criticallity and technicallity		**R4** Limit RFI review loops and perform joint-reviews for RFIs exceeding the number of communication loops
	R5 Set autorejection worflows for Incomplete or non-conformant RFIs		**R5** Periodic automated process performance monitoring (cost efficiency)
	R6 Establish a clear RFI routing procol & Standard Operating Procedures (SOP)		**R6** Automated detection & resolution of procurement-related RFIs
	R7 Frequent process performance mornitoring with automated alerts		**R7** Prioritize process redesign of high-cost variants for future similar RFIs
	R8 Orchestrate the RFI process (routing RFIs) using AI agents		**R8** Conduct constructability reviews and design completeness reviews early in the project
	R9 Predifine decision-making criteria for the approval or confirmation of Value Engineering or Design alternatives		**R9** Define standard RFI template and mandatory fields flagging cost impact in Standard Operating Procedures (SOP)
	R10 Monitor and control weekly interraval rates of RFIs (promote front-loaded submissions & reviews but wathc for abusive amount/misuse of RFIs)		**R10** RFI Cost forecasting for future similar projects based on assessment results
			R11 AI-Assisted duplicate detection of RFIs based on content similarity
Organizational (O)	**R1** Define RFI escalation protocols by role and specialization	**Organizational (O)**	**R6** Periodic automated performance monitoring & reporting by team roles against organizational benchmarks
	R2 Delegate management activities to project coordinators		**R7** Automated detection of overloaded resources & workload balancing by roles
	R3 Specify RACI matrix by role in Standard Operating Procedures (SOP)		**R8** Implement time-bound joint RFI reviews of roles frequently involved in high-delays
	R4 Define cross-functional reviews for high-technical or critical RFIs		**R9** Predict 'best-fit role' to review RFI based on RFI conent
	R5 Provide staff training to efficiently manage RFIs within the CDE		**R10** Implement dynamic assignment of RFI reviews considering RFI criticality, technical complexity, and staff availability: automated evidence-based RFI re-routing.

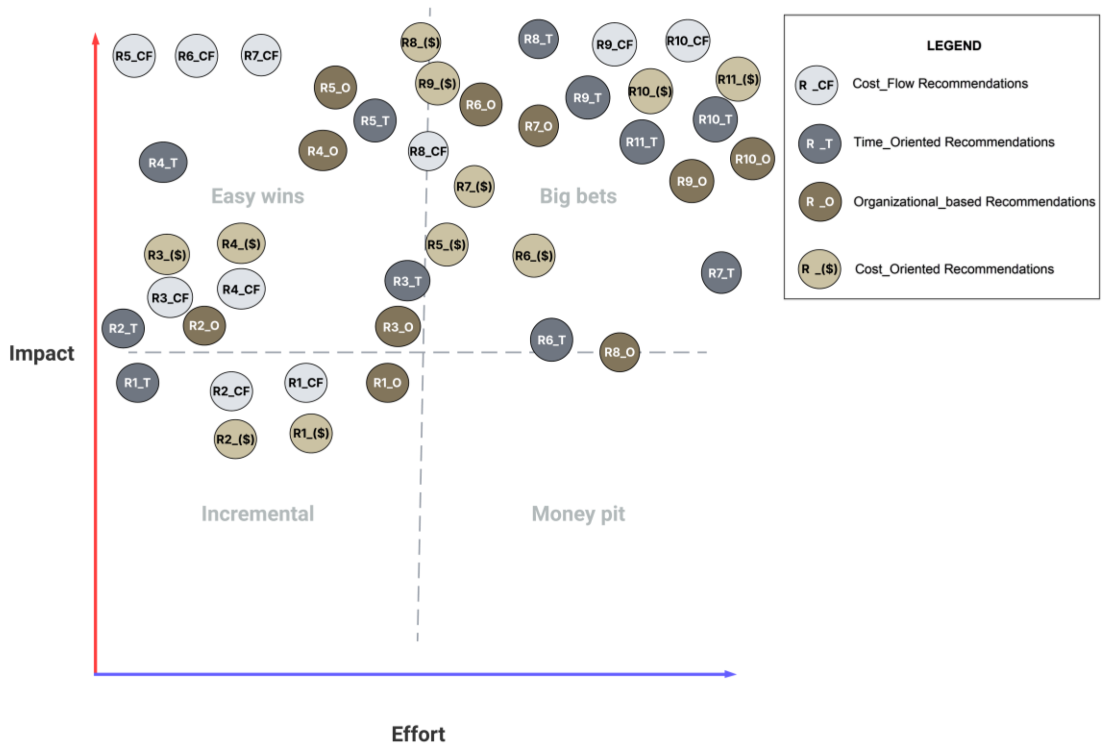


Figure 6-23. Prioritized RFI process improvement recommendations derived from data-driven analysis.

The prioritization presented in Figure 6-23 was performed using the following criteria. For the impact scale: minimal impact (addresses rare or low-frequency deviations), moderate impact (addresses mid-range bottlenecks), and high impact (addresses dominant process inefficiencies observed in frequent variants). For the effort scale: minimal effort (involving parameter adjustments and the establishment of target thresholds), medium effort (requiring new SOPs, limited training, and minor system customization), and high effort (entailing governance changes, redefinition of contract-level agreements, and AI-powered targeted automations). The discussion section of this study provides further details on these recommendations, with examples for each process perspective.

6.2.3.2. Module 7 (M7): Continuous Process Performance Monitoring & Dynamic Reporting

Although the implementation of the derived process improvement recommendations lies outside the scope of this study, successful continuous improvement requires periodic automated process performance monitoring with dynamic reporting. This ensures that the reengineering process remains targeted and action-oriented, rather than resulting in costly automation efforts that fail to deliver value when blindly undertaken without proper assessment, monitoring, and reporting. An example of the dynamic performance monitoring reports that can be created and enriched for the RFI process of P044 is presented in Figure 6-24. This real-world dynamic process performance report was designed based on the designed operational process management canvas introduced in Section 3.5.1 to audit and monitor various operational aspects related to the RFI management.

Such dynamic reports can be generated using process mining software [97, 102, 105, 266] following the implementation of the proposed process-aware LPMM methodology. They can be configured to reflect ‘as-happened,’ semi–real-time, or real-time analyses, depending on the established backend process-aware architecture and defined data pipelines. The report in Figure 6-24 highlights the most frequent ‘as-happened’ process execution and presents various KPIs, including the number of RFIs, the total number of activities following the most common path (505), the total number of distinct process activities (75), the total number of RFI events (1660), the average RFI process lead time, detected deviations and their potential root causes, key variations, bottlenecks and violations with negative impacts, as well as the visual throughput time of organizational roles.

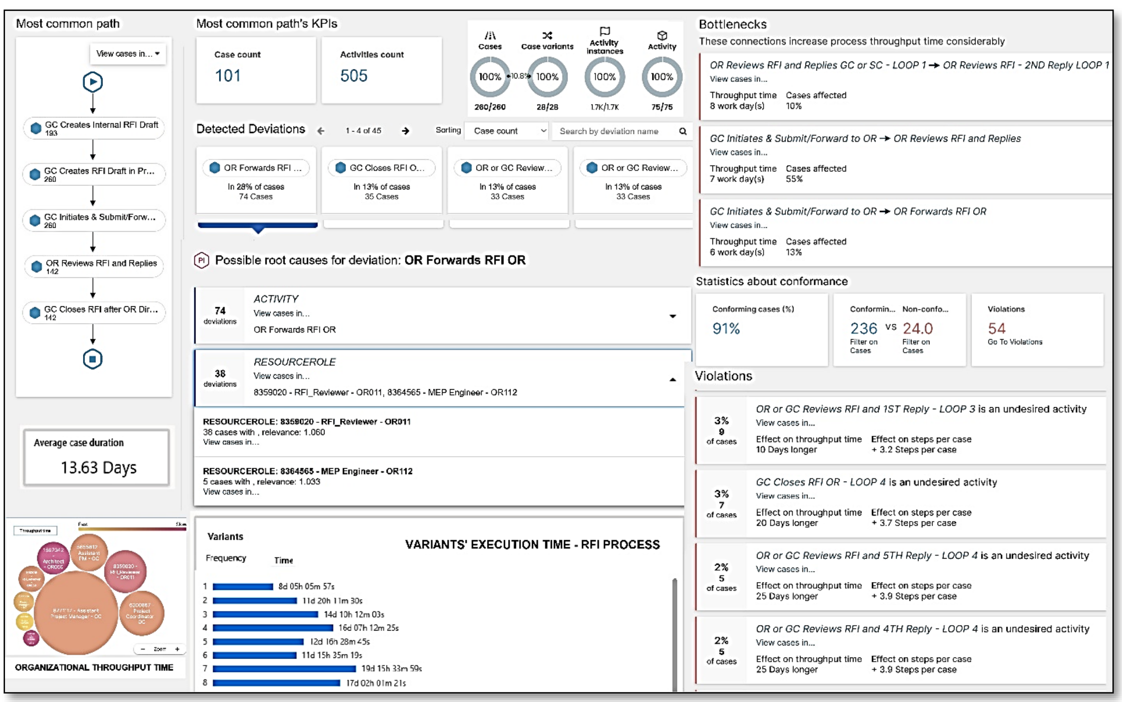


Figure 6-24. Dynamic reporting on automated RFI process performance monitoring (P044)

6.3. Discussion

In construction, significant emphasis is placed on 'project' controls to ensure delivery on time, within budget, and with managed risks. Yet, automated and data-driven 'process' control, along with PHM of real-world construction business operations remain largely unaddressed, leaving projects exposed to hidden risks of delays and cost overruns. In this vein, the proposed LPMM approach in this study seeks to bridge this gap by providing a systematic data-driven method to enable the automated extraction of a digital process model of the actual RFI management workflow executions, while enabling continuous assessment and dynamic monitoring of the RFI process performance through quantitative, automated, and comprehensive cross-case analysis across four essential perspectives.

The control-flow analysis revealed that the 'as-happened' RFI process in P044 was highly fragmented, characterized by numerous variants (28) and detours (8 loops). The dominant process comprised 32 frequently executed activities out of 75, together with multiple back-and-forth loops that resulted in substantial delays of up to 61 days and additional rework. Deviations included skipped or unanswered RFIs and lengthy multi-stage reviews, particularly in variants with multiple forwarding and value-engineering (VE) decisions. These inefficiencies were amplified by excessive communication loops, batching practices, and the involvement of numerous approvers, which collectively formed major bottlenecks. To address these challenges, targeted and action-oriented recommendations were derived from the data-driven process assessment and diagnosis. These include establishing thresholds for communication loops (i.e., < 4), defining clear routing protocols and decision-making criteria, prioritizing RFIs based on criticality and technicality, predefining reviewing authorities per RFI type, and automating rejection workflows for incomplete submissions. Additional measures such as frequent performance monitoring with alerts, orchestration of routing through AI agents, and stricter control of submission intervals are proposed to avoid batching periods, enhance transparency, streamline decision-making, and improve overall process performance.

The time-oriented analysis unveiled various inefficiencies including prolonged lead times (> 3 weeks) influenced by infrequent executions, ping-pong communication patterns, batching practices, and process drifts during review stages. P044 presented a low weighted average RFI process efficiency (0.33). In P044, RFI work peaks occurred in the middle and toward the end of the process, reflecting a reactive and corrective approach, whereas in P063, peaks occurred at the beginning, indicating a more proactive and preventive strategy. Comparative results showed that P063 was less reactive, with fewer extended reviews, shorter communication loops, and reduced process variations, leading to improved timeliness. To address these issues, the study recommends establishing SLA-based resolution deadlines tailored to RFI types, enabling early identification of RFIs requiring on-site verification, and implementing workload balancing and fast-track review workflows for critical RFIs. Additional measures include adopting the proposed LPMM methodology grounded in Lean and Agile metrics for continuous monitoring, forming task-force reviews with only essential parties, deploying automated alerts to detect deadline risks and activity drifts, and implementing smart reviewer assignments based on competency and workload. These actions aim to minimize time losses, enhance responsiveness, and improve the predictability of RFI process performance.

The organizational analysis highlighted imbalances in resource allocation, accountability gaps, and workload concentration across specific RFI reviewers, particularly the Assistant PM GC (ID = 8771117), who emerged as a central but overloaded resource contributing to extended lead times. Discovered patterns of accountability "ping-pong" and multiple forwardings further fragmented responsibilities, while certain roles such as Architects, OR Reviewers, and Mechanical Engineers were consistently associated with prolonged reviews. To mitigate these inefficiencies, the recommended actions emphasize clarifying responsibilities through escalation protocols and Responsibility, Accountability, Consulted, and Informed (RACI) matrices established early in the project, delegating routine management activities to coordinators, and strengthening cross-functional reviews for technical RFIs. Additional measures include automated monitoring of role workloads, predictive assignment of the most suitable reviewers, and dynamic role allocation based on availability, technical expertise, and backlog. The main goal from this perspective is to streamline accountability, balance workloads, and enhance organizational responsiveness in the RFI process.

Finally, the cost-oriented assessment highlighted several inefficiencies, particularly in Project P044, which incurred higher costs and lower efficiency (WACE = 0.24) compared to Project P063 (WACE = 0.67). Redundant submissions, unclear role responsibilities, and excessive review loops contributed to substantial cost overruns. Mitigation strategies under this perspective include tiered and escalated reviews, cost-sharing agreements, automated monitoring, AI-based duplicate detection, and standardized templates. These measures are intended to reduce accrued costs, increase overall efficiency, and enhance cost predictability and control across RFI process executions. For owners, these recommended actions provide greater oversight and help prevent the abusive use of RFIs, while for general contractors, they aim to increase productivity and minimize work disruptions through faster turnaround times.

6.4. Concluding Remarks

In contrast to existing approaches, the proposed systematic LPMM approach presented in this study introduces a novel, data-driven, quantitative method for the automated and dynamic assessment and monitoring of the actual execution performance of intra- and inter-organizational construction business processes within the bounds of the captured digital footprints. The proposed method was applied to real-world settings through the developed RFI process mining use case. The soundness of the automated RFI process model extraction module was verified using process model discovery metrics, which demonstrated satisfactory quality performance in terms of fitness (0.958), precision (0.761), and F1-score (0.845). Moreover, the proposed approach for dynamic, quantitative process performance assessment and monitoring was validated through a real-world cross-case analysis. This analysis comprised the automated discovery and quantitative performance evaluation of 5,564 E2E RFI process management executions across 71 actual construction projects, all of which were managed by a GC as the lead party appointed by its clients. The results indicate that the process efficiency of RFI management can improve by up to threefold through enhanced process control, generating cost savings of approximately $114,000 USD in a single large commercial project (P044). When extrapolated across project portfolios and their underlying inter-organizational processes, these improvements reveal substantial organization-wide and industry-wide opportunities, underscoring the transformative impact of data-driven process improvement in enhancing resource allocation, strengthening operational control, and achieving scalable performance gains.

This study advances the field of data-driven process management and control in construction projects, offering fact-based, actionable process-oriented insights for project directors, project and innovation managers, project coordinators, and project controls specialists. The main contributions and benefits of the study are summarized as follows:

(1) *Full ETL+E Pipeline for Automated Process Event Log Generation* – This entailed the algorithm development of (i) an extractor algorithm and automated python functions for the automated retrieval of project operations data from the Procore CDE powered by a configured API-based data connection application, (ii) a transformer algorithm that converts raw system data into a process event-log structure capturing the process context, and (iii) an enricher algorithm that integrates the cost dimension into the event logs, combined with an XES serializer and logger, thereby generating standard-compliant logs that enable robust process mining implementations and strengthen the corporate knowledge memory of construction organizations' business operations.

(2) *Automated extraction of "as-happen" process models to support data-driven process performance assessment and process improvement in RFI and Construction Change Order management* – This included (i) the development of the process mining-based use cases for these processes, (ii) the automated extraction of the 'as-happened' E2E process models, (iii) the quantitative assessment of process performance using real-world project operations data, comprising the RFI process efficiency evaluation at the GC's portfolio level and cross-case analysis at the process level across four essential dimensions (time, cost, control-flow, and organizational) to detect main process performance inefficiencies, enhance resource allocation, reduce costs, and lower decision-making risks, (iv) dynamic Process Health Monitoring reporting to increase visibility and transparency of process performance, and (v) actionable process improvement recommendations for re-engineering, streamlining, and enhancing these processes informed by fact-based performance diagnostics.

These contributions, together with acknowledged limitations and future research directions, are consolidated and synthesized in Chapter 7. , in alignment with the overarching research objectives.

CHAPTER 7. RESEARCH DISCUSSIONS

To address the control-flow inefficiencies identified in Section 6.2.2.2 and summarized in Table 6.5, several targeted interventions are recommended. These include establishing thresholds for communication loops (e.g., <4 iterations), defining clear routing protocols and decision criteria, prioritizing RFIs based on criticality and technical complexity, predefining reviewing authorities by RFI type, and automating rejection workflows for incomplete submissions. Additional measures include continuous performance monitoring with automated alerts, AI-assisted routing orchestration, and stricter control of submission intervals to prevent batching periods. Together, these actions aim to enhance transparency, streamline decision-making, and improve overall process performance. Further improvements under this perspective include implementing SLA-based resolution deadlines tailored to RFI types, enabling early identification of RFIs requiring on-site verification, and deploying workload-balancing and fast-track review workflows for critical RFIs. Other recommended measures involve adopting the proposed D2P method for continuous data-driven diagnostic process performance monitoring, organizing focused task-force reviews involving only essential stakeholders, deploying automated alerts to detect deadline risks and process drifts, and implementing intelligent reviewer assignments based on competency and workload. Collectively, these actions aim to reduce delays, enhance responsiveness, and improve the predictability of RFI process performance.

Recent work by Yilmaz et al. (2025) trained machine-learning models to predict RFI closure times using operational data from a large airport project [267]. Although their study demonstrated predictive capability, the E2E process explainability of the models remained limited, as the approach relied on an “as-planned” RFI process representation and lacked data-driven diagnostic analytics capable of revealing actual execution inefficiencies. In contrast, the present study analyzes the real “as-happened” process behavior derived from event logs. Some factors influencing RFI duration prediction were identified in [267], including technical complexity, RFI content types, and discipline-specific complexity, were also observed in our analysis. These factors are integrated and augmented with the evidence-based outcomes derived from the diagnostic performance analysis conducted in this study, thereby improving predictive model explainability and potentially enhancing prediction accuracy. Moreover, [267] recognized that RFI inefficiencies propagate into downstream project control processes [26]. which aligns with the findings of the present study. A further distinction concerns observed RFI closure durations. In the two large commercial projects analyzed here (P044 and P063), the average RFI closure time remained within 14 days, although some RFIs deviated significantly, reaching over 70 days in P044 and 45 days in P063. By contrast, the large-size airport project examined in [267] reported an overall average closure time of approximately 50 days. These differences highlight the influence of RFI technical complexity, RFI content types, and contextual project characteristics on resolution timelines.

At the organizational level (see Section 6.2.2.2 and Table 6-5) mitigation strategies focus on clarifying responsibilities through escalation protocols and Responsibility-Accountability-Consulted-Informed (RACI) matrices established early in the project lifecycle. Additional actions include delegating routine coordination tasks to dedicated project coordinators and strengthening cross-functional reviews for technically complex RFIs. Supporting measures include implementing standard operating procedures or integrated framework protocols as proposed by Golzarpoor et al. (2016) [24, 169], automated monitoring of role workloads, predictive assignment of appropriate reviewers, and dynamic role allocation based on availability, technical expertise, and backlog conditions. These measures aim to streamline accountability, balance workloads to avoid overloaded resources that create process bottlenecks [90] and improve inter-organizational responsiveness within the RFI process.

From a cost-oriented perspective, mitigation strategies include tiered and escalated review structures, cost-sharing agreements, automated monitoring of RFI processing stages, AI-based detection of duplicate RFIs, and the use of standardized submission templates. These measures aim to reduce accrued costs, improve efficiency, and enhance cost predictability across RFI process executions. For owners, they strengthen oversight and reduce the potential misuse of RFIs, while for general contractors they mitigate productivity losses by minimizing fieldwork disruptions through faster review cycles.

Despite the growing emphasis on process improvement and digital transformation in construction, many RFI process improvement initiatives [242, 243] tend to overlook the prerequisite of discovering and quantitatively analyzing the performance real-world E2E process executions to diagnose process inefficiencies. Without such evidence-based insights, process improvement decisions remain largely intuitive and disconnected from real-world business operations. Existing studies within the AEC/FM literature provide useful cost-based benchmarks and industry best practices [53, 227, 243, 251, 264] but they are typically static and time-bound, limiting their ability to capture E2E execution-level, cost-based process variability and diagnose root causes of suboptimal performance in RFI management.

The proposed LPMM-D2P method directly addresses several key challenges to process mining implementation identified by Zimmerman et al. (2024) by systematically bridging the gap between data-oriented project environments and process-aware analytical structures [268]. In particular, the method mitigates data extraction and integration limitations (C4, C6) through automated API-based extractors and a system-agnostic integration approach. Challenges related to data structure knowledge and transformation (C7, C8) are addressed through the formalized D2P transformation pipeline, which defines standardized event log schemas compliant with XES. Data quality and validation issues (C9, C10) are explicitly handled through structured preprocessing, enrichment, and validation steps within the pipeline. Furthermore, the method enhances domain understanding and question formulation (C1, C20) via the use of process-oriented use case definitions and expert-informed aggregated inputs through semi-structured interviews [46]. Finally, by enabling multi-perspective analysis and continuous process health monitoring, the approach supports improved analysis focus, interpretation, and decision-making, thereby reducing barriers associated with process mining adoption in complex construction project environments.

The proposed LPMM–D2P method builds upon previously established process mining methodologies such as PM², but it differs fundamentally from them by explicitly addressing the data-to-process transformation challenge that is typically assumed rather than formalized. While PM² provides a generic lifecycle for conducting process mining analyses, it presumes the availability of well-structured event logs and offers limited guidance on extracting and transforming heterogeneous data from complex information systems. In contrast, the D2P method introduces a system-agnostic, automated pipeline for extracting, structuring, and enriching project data into process-aware event logs compliant with XES. Furthermore, LPMM extends beyond traditional analysis by integrating multi-perspective performance assessment, including control-flow, organizational, and cost dimensions, as well as enabling continuous Process Health Monitoring (PHM). In addition, the method incorporates Lean-based performance metrics, such as value-added time and process efficiency, to support waste identification and continuous improvement in construction workflows. The proposed LPMM–D2P framework is specifically tailored to the construction domain, accounting for the complexity of inter-organizational processes and heterogeneous project information systems. This combination of data architecture transformation, domain-specific adaptation to construction processes, and continuous monitoring capabilities positions the proposed method as a novel extension that bridges the gap between raw project data environments and actionable process intelligence, which is not explicitly addressed in prior methodologies such as PM² [210].

The D2P method implemented in this study addresses this gap by enabling systematic extraction and quantitative analysis of "as-happened" process behaviors from heterogeneous, siloed project data sources. Without such diagnostic insights, process reengineering and automation initiatives risk being poorly targeted and ineffective. In contrast, D2P enables evidence-based interventions, such as intelligent dynamic assignment of RFI reviewers and automated rerouting of technically complex RFIs to expert consultants, thereby reducing unnecessary activities such as repeated forwarding and review loops (see Table 6-6)These capabilities support more effective control of process variability and enable targeted improvements together with enhanced decision-making grounded in data-driven process analytics.

CHAPTER 8. RESEARCH CONCLUSIONS AND OUTLOOK

8.1. Summary

This study introduces a novel data-driven, process-oriented management framework tailored to the AEC/FM domain to drive process intelligence in construction projects. The proposed process-aware framework presents a Lean-based Process Mining and Management (LPMM) method that integrates advanced process mining techniques with Lean-based quantitative metrics to help construction organizations and practitioners leverage day-to-day project operations data to automatically and dynamically assess, control, and manage construction business operations. It should be noted that the LPMM method is system-agnostic (with respect to software platforms and data standards), enabling the automated extraction and integration of fragmented project operations data residing across multiple digital construction platforms widely adopted across the industry, such as Procore and other comparable CDEs, to support E2E process performance monitoring.

8.2. Contributions

8.2.1. Contributions – Research in AEC/FM

8.2.1.1. Fundamental Contributions

(1) ***<u>Lean-based Process Mining and Management (LPMM) framework (Primary)</u>*** – The primary contribution of this process-oriented research is the introduction of a novel Lean-based Process Mining and Management (LPMM) framework tailored to the AEC/FM domain. The proposed framework enables automated extraction of *'as-happened'* process models, advanced multi-perspective performance assessment, and continuous monitoring of critical business operations in construction projects. To the best of the author's knowledge, no prior framework has provided an integrated, Lean-grounded, process-mining-based methodology addressing this need in the AEC/FM industry (see Figure 5-1). The complementary contributions associated with this framework are as follows:

a. *<u>Identification and validation of significant construction business processes:</u>* Process significance was systematically validated through semi-structured interviews with industry experts and supported by data-driven analytics, including exploratory data analysis (EDA) and task mining techniques (see Chapter 4.

b. *RFI Process Mining Use Case Development* – A comprehensive RFI process mining use case was developed, comprising: (i) in-depth process understanding based on an extended literature review; (ii) definition of scope for quantitative process performance assessment; (iii) identification of the primary information systems supporting RFI execution, stakeholder roles, core data entities, and human–machine interactions (HMI); (iv) formulation of a backend data processing pipeline; and (v) multi-perspective process-oriented analytics incorporating time, cost, control-flow, and organizational dimensions (see Figure 6-5).

c. *Design of reusable D2P transformation method*: The proposed method is replicable to other significant construction business processes beyond RFIs (see Figure 3-7).

(2) *Enablement of RFI Process Health Monitoring (PHM) –* A second primary contribution of this study is the development of a Process Health Monitoring (PHM) module designed to operationalize continuous, evidence-based process performance assessment and monitoring. The PHM module comprises three integrated components: (i) identification of key Process Performance Indicators (PPIs) grounded in the existing Process Mining and Management (PMM) literature; (ii) formulation of time-based and cost-based process efficiency equations (see (Eq. 9) and (Eq. 10)), derived from Lean-based performance metrics and enriched through process mining-based variant analysis; and (iii) introduction of an operational business process model canvas and a dynamic process board reporting mechanism to support structured process governance (i.e., formalized oversight, control, and continuous improvement of business processes) and decision-making. The complementary contributions include:

a. *Cost-based Process Variant Analysis –* Automatically discovered process variants were enriched through the integration of a cost dimension into the generated event logs, computed by multiplying activity cycle times by stakeholder-specific hourly wage rates. This extension enabled quantitative evaluation of variant-level economic performance and supported cost-informed process comparison. It further allowed AEC/FM organizations to identify dominant and non-conformant process variants, thereby facilitating early, targeted interventions to mitigate cost inefficiencies and enhance process performance.

b. *Definition of Time-based and Cost-based Process Efficiency Metrics* – Formulation of mathematically defined time-based and cost-based efficiency equations (see (Eq. 9) and (Eq. 10)), derived from Lean performance principles and process mining-based variant analysis. These metrics provide a structured mechanism to assess operational efficiency across alternative process executions.

c. *Operational Business Process Model Canvas:* Design of a structured seven-component operational process canvas (see Figure 3-8) integrating automation level, process purpose and abstraction, information structure, maturity assessment, monitoring mechanisms, classification, and model representation. The canvas provides a standardized blueprint for systematic process analysis, audit, and performance monitoring.

8.2.1.2. Applied Contributions

(1) *Design and Implementation of an automated D2P transformation pipeline –* This research designed and implemented an automated Data-to-Process (D2P) transformation pipeline that systematically converts heterogeneous raw system data into structured process knowledge assets suitable (i.e., in conformance with XES standard) for process mining and performance analyses. The proposed pipeline transforms fragmented digital trace data into analyzable process knowledge (data as an asset). This research designed and implemented an automated Data-to-Process (D2P) transformation pipeline that systematically converts heterogeneous raw system data into structured process knowledge assets compliant with the XES standard for process mining and performance analysis. The proposed pipeline transforms fragmented digital trace data into analyzable, structured process events, thereby operationalizing data as a strategic asset for evidence-based performance monitoring.

(2) *Evidence-based performance assessment enabled through LPMM –* This study operationalized the LPMM framework to enable a multi-perspective, evidence-based process performance assessment across four analytical perspectives: (i) temporal

performance (cycle times, waiting times, and throughput), (ii) cost-based process variant analysis (delay exposure and variant-level financial impact), (iii) control-flow variability and conformance relative to a dominant or reference process model, and (iv) organizational interaction patterns using SNA to examine how stakeholder roles and team dynamics/interaction structures influence overall actual E2E process performance as well as to enhance resource allocation (i.e., labor, materials, machinery, money) by automated resource balancing based on detected workloads (i.e., WIP) by similar project roles.

(3) *Multi-project cross-case validation of the proposed E2E RFI process performance monitoring module –* This contribution empirically validated the proposed structured monitoring module across multiple projects by integrating automated D2P transformation, RFI process mining-based analytics, and the defined PPIs. The results demonstrated the methodological feasibility and cross-project applicability of systematic E2E RFI process performance monitoring by leveraging project operations data that is frequently underutilized or remains dormant within cloud-based CDEs. In parallel, process custom boards by perspective (see Figure 6-24) enabled continuous visualization, analysis, and tracking of key defined PPIs and variant-level metrics, thereby supporting structured and data-driven performance monitoring.

8.2.2. Contributions to AEC/FM industry – Development and Deployment

8.2.2.1. Fundamental Contributions

(1) *RFI Management as a Diagnostic Process System –* While prior studies have examined RFI management from a process-oriented perspective, they primarily relied on conceptual models, reference frameworks, or survey-based assessments of intended workflows. This research advances the field by repositioning RFI handling from a document-centric administrative workflow to a measurable, data-driven inter-organizational business process grounded in automatically extracted "as-happened" project operations data. By structuring RFIs as event-based process objects and analyzing actual execution traces rather than assumed process flows, the study enables systematic lifecycle analysis and diagnostic performance assessment, allowing the empirical identification and quantification of inefficiencies such as delays, rework loops, and process variability across projects. The

contribution emphasizes evidence-based performance diagnostics of real process executions, addressing an important industry need that had received limited attention.

(2) *Development of ETL+E Algorithms for Automated Process Event Log Generation –* It comprised the development and deployment of automated extractor, transformer, and enricher algorithms to retrieve project operations data from Procore CDE, convert raw system data into structured event logs, and integrate cost dimensions, generating standard-compliant XES logs that enable robust process mining and enhance organizational operational knowledge retention (see Sections 5.4.1.3 and 6.4).

8.2.2.2. Applied Contributions

(1) ***Real-World Deployment and Multi-Project Validation of the D2P Pipeline (Primary) –*** This contribution entailed the full operational validation and deployment of the D2P pipeline using real RFI data extracted from production CDEs and SQL databases across multiple construction projects. The validation demonstrated methodological robustness and feasibility under realistic data complexity and volume conditions. The deployed pipeline supported scalable cross-project and portfolio-level process analytics, generating actionable, data-driven insights tailored to project managers, coordinators, and controls specialists to enhance process transparency, responsiveness, and informed decision-making. This empowers AEC/FM organizations to evolve their existing data infrastructures into process-aware environments that enable actionable performance diagnostics and support future developments in evidence-based predictive capabilities.

8.3. Attainment of Research Objectives and Questions

This research project aims to drive process intelligence in construction projects by offering a data-driven method based on process mining and Lean-based metrics, helping AEC/FM organizations identify major process inefficiencies automatically and in near real time, and enabling project stakeholders to take immediate action when needed to improve efficiency and reduce cost overruns.

The five research objectives defined in this study have been fully achieved as follows:

- RO1 was accomplished through an investigation of the foundational principles of

Process Mining and Process Modeling and Management (PMM), including a comprehensive review of international event-log standards, PMM methodologies, their research gaps, and the identification of construction service-based administrative processes required across the lifecycle of construction projects. This also included the proposal of a process management support tool for conducting process audits and monitoring and controlling these processes.

- RO2 was achieved by validating the identified administrative processes and their significance in project delivery through semi-structured interviews with subject matter experts to capture tacit process knowledge and corroborate the literature review findings. In addition, operational project data were collected, and a preliminary exploratory and task mining analysis was performed to select the most relevant process or processes for subsequent process mining analysis.
- RO3 was fulfilled by designing the process mining use cases for the selected RFI Management and Change Order Management processes, which resulted in a robust end-to-end understanding of their execution and interdependencies.
- RO4 was addressed by developing the comprehensive LPMM methodological framework, including several supporting algorithms for automated event-log generation required for process mining implementation.
- RO5 was completed by applying and validating the proposed method in actual construction projects through real-world case studies, demonstrating the practicality, applicability, and usefulness of the framework.

Besides, the following two main research questions were also addressed:

- RQ1: How can construction organizations harness project operations data to extend and transform their existing data-oriented architectures into process-aware structures to control the performance of their business operations?
    - ➢ This research demonstrated that implementing the PMM methodological framework presented in Figure 5-1 enables construction organizations to systematically and automatically transform their existing data-oriented

architectures into process-aware structures, thereby leveraging project data generated from day-to-day operations to continuously assess, monitor, and improve the execution performance of their intra- and inter-organizational business processes.

- RQ2. How can construction professionals leverage process mining, in combination with other process-oriented methodologies, extract, quantitatively assess, monitor, and improve the actual E2E performance of critical construction processes in a dynamic, automated, and data-driven manner?
- The proposed process mining-based framework was integrated with time- and cost-efficiency equations (see (Eq. 9) and (Eq. 10)) derived based on Lean-based metrics, analyzing the impact of non-value adding activities on process efficiency.

8.4. Limitations and Future Work

Despite the contributions of this study and the achieved objectives, the following limitations are acknowledged: (i) although the proposed approach is designed to be system-agnostic and generally applicable across different Common Data Environments (CDEs), its empirical implementation and validation were conducted using data obtained from SQL-based relational databases and a single CDE platform, Procore, through the automated D2P transformation method. The selection of this CDE reflects its widespread adoption among AEC/FM organizations in North America [19, 234]; (ii) Process simulation and predictive analytics are not addressed in this study. Instead, the study adopts a diagnostic process analytics perspective, aiming to identify and quantitatively diagnose major performance inefficiencies in real-world RFI process executions; and (iii) the analyzed GC's operations data is confined to DBB construction projects.

In line with the notion of 'No AI without PI' [269], this study underscores automated process performance assessment and diagnostics as the essential foundation for enabling targeted, non-wasteful AI-powered automations, reinforcing the principle that what cannot be assessed cannot be controlled. It proposes data-driven process improvements through the LPMM framework as the enabler of generative, predictive, and prescriptive AI, thereby streamlining workflows and enhancing decision-making, while the actual implementation of AI-powered automations lies beyond its scope. Future research may extend this process-oriented study and advance the D2P framework by: (i) examining the performance implications of project delivery systems beyond DBB; (ii) developing additional automated data pipelines to other CDEs integrated with BIM 3D models; (iii) incorporating multidimensional process mining based on the OCEL standard, enabling data-driven analysis of interdependent process executions and the systematic assessment of intertwinement effects across interconnected construction business processes, e.g., E2E change order management and its interactions with RFIs, change order field execution activities, and progress payment processes.; (iv) extending the current framework toward predictive, data-driven, process-oriented analytics with forward-looking capabilities; (v) assessing performance improvements under "what-if" scenarios through the simulation and replay of generated process event logs; and (vi) designing targeted AI-powered agentic automations grounded in evidence-based decision-making and process variant control strategies to streamline construction business operations through integrated diagnostic and predictive process intelligence. This study, together with the identified future directions, sets the foundation for more predictable, efficient, and digitally enabled project delivery by advancing data-driven process performance assessment, diagnostics, and targeted continuous improvements, with a particular focus on addressing the most critical process inefficiencies.

8.5. Threats to Research Validity

Construct Validity – Event Logs Representation and Modeling Assumptions: The transformation of raw system data into event logs relies on modeling assumptions regarding case identifiers, activity definitions, and lifecycle transitions, which may influence the extent to which the logs accurately represent real process behavior.

External Validity – Generalizability: The method is validated on the RFI process, which is inherently interconnected with other construction processes such as design authoring, contract management, and financial workflows, and therefore further studies are required to confirm its generalizability across broader intertwined/interrelated process contexts.

Conclusion Validity – Limited Statistical Generalization: Although the study includes 71 selected projects, the analysis is restricted to a single process type and does not incorporate formal statistical analysis or statistical process control methods. The methodological approach emphasizes process variant analysis to uncover execution patterns, deviations, and inefficiencies across cases. While this provides rich diagnostic insights into real-world process behavior, it limits the extent to which the findings can be generalized in a statistical sense.

Data Quality, Completeness, Availability, and Privacy: The validity of the analysis depends on the availability, completeness, and accuracy of digital project data, while data access and privacy constraints may further limit the scope of observable process behavior.

Researcher Bias: The reconstruction of event log structures, process event log abstraction, and performance metrics may introduce researcher bias that can influence the interpretation of process models and analytical performance results.

8.6. Synthesis of Research Impact

This research strengthens the operational knowledge base of AEC organizations (i.e., corporate memory) through the implementation of the proposed Data-to-Process (D2P) method, which leverages process mining and quantitative Lean-based performance metrics (LPMM) to systematically assess, diagnose, and monitor the performance of essential construction business processes. This hybrid and structured methodology enables the automated extraction of "as-happened" process executions, followed by diagnostic analysis to identify major inefficiencies and performance bottlenecks where timely and targeted interventions are required. In doing so, it enhances process transparency, increases visibility over actual process behavior, enables continuous monitoring and operational control of process variants, and supports long-term productivity improvement [26].

Furthermore, this research provides a systematic pathway for integrating fragmented, siloed data systems and transforming existing data-oriented architectures within AEC/FM organizations into process-aware structures. By operationalizing data management and process variant control strategies, it ultimately drives process intelligence and enhances process efficiency through evidence-based decision-making and targeted automations.

In conclusion, the outcomes of this research equip project managers, project controls specialists, and coordinators with systematic, data-driven capabilities to monitor, diagnose, and control operational performance more effectively. By enabling the automated identification, evaluation, and diagnosis of process inefficiencies, the proposed approach supports timely and targeted early interventions, shifting decision-making from reactive problem resolution to proactive performance monitoring and control. This transition toward anticipatory management of process variants enhances operational efficiency, strengthens productivity, and ultimately contributes to measurable time and cost savings across construction projects.

REFERENCES

[1] McKinsey, "Reinventing construction through a productivity revolution," *McKinsey&Company,* 2017. [Online]. Available: https://www.mckinsey.com/business-functions/operations/our-insights/reinventing-construction-through-a-productivity-revolution.

[2] A. J. Martinez Lagunas and M. Nik-Bakht, "Analysis of XES and OCEL Data Schemas: Towards Multidimensional Process Mining of Intertwined Construction Processes," (in English), *Proceedings of the 41st International Symposium on Automation and Robotics in Construction,* vol. Construction Management Techniques, no. 2413-5844, pp. 444-451, 2024, doi: https://doi.org/10.22260/ISARC2024/0058.

[3] McKinsey Global Institute, "Digital America: A tale of the haves and have-mores.," 2015. Accessed: September 01, 2025. [Online]. Available: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/digital-america-a-tale-of-the-haves-and-have-mores

[4] McKinsey&Company, "Reinventing Construction: A route to higher productivity," 2017. [Online]. Available: https://www.mckinsey.com/capabilities/operations/our-insights/reinventing-construction-through-a-productivity-revolution

[5] L. Espina-Romero *et al.*, "Which Industrial Sectors Are Affected by Artificial Intelligence? A Bibliometric Analysis of Trends and Perspectives," *Sustainability (Switzerland),* vol. 15, 2023, doi: https://doi.org/10.3390/su151612176.

[6] F. Jalaei, Y. Sultan, E. A. Poirier, S. Staub-French, D. Bhonde, and S. Davari, "Digitalization and Productivity Research and Development Roadmap," pp. 1-68, 2024, doi: https://doi.org/10.4224/40003328.

[7] M. Malinova, S. Gross, and J. Mendling, "A study into the contingencies of process improvement methods," *Information Systems,* vol. 104, 2022, doi: https://doi.org/10.1016/j.is.2021.101880.

[8] Deloitte, "Global Deloitte Process Mining Study 2025," 2025. Accessed: March 21, 2025. [Online]. Available: https://www.deloitte.com/de/de/services/financial-advisory/research/global-process-mining-survey.html

[9] HSPI Management Consulting, "Process Mining: A Database of Applications," Version 2.0 2021. Accessed: June 15, 2025. [Online]. Available: https://www.tf-pm.org/news/a-2021-edition-of-the-extensive-list-of-process-mining-applications-has-been-released-by-hspi-s-p-a

[10] M. M. Queiroz, S. Fosso Wamba, M. C. Machado, and R. Telles, "Smart production systems drivers for business process management improvement: An integrative framework," *Business Process Management Journal,* vol. 26, no. 5, pp. 1075-1092, 2020, doi: https://doi.org/10.1108/BPMJ-03-2019-0134.

[11] W. M. P. Van der Aalst, *Process Mining Data Science in Action*, 2nd ed. Berlin, Heidelberg: Springer (in en), 2016.

[12] M. Weske, *Business Process Management*, Second Edi ed. Springer, 2013, pp. 1689-1699.

[13] M. Dumas *et al.*, "AI-augmented Business Process Management Systems: A Research Manifesto," *ACM Transactions on Management Information Systems,* vol. 14, no. 1, 2023, doi: https://doi.org/10.1145/3576047.

[14] M. M. Queiroz, S. C. F. Pereira, R. Telles, and M. C. Machado, "Industry 4.0 and digital supply chain capabilities: A framework for understanding digitalisation challenges and opportunities," *Benchmarking,* 2019, doi: 10.1108/BIJ-12-2018-0435.

[15] Deloitte, "State of Digital Adoption in the Construction Industry 2025 Autodesk," 2025. Accessed: 2025-09-01. [Online]. Available: https://www.deloitte.com/au/en/services/economics/analysis/state-digital-adoption-construction-industry.html

[16] A. Sawhney, M. Riley, J. Irizarry, and C. T. Pérez, "A proposed framework for Construction 4.0 based on a review of literature," 2020, vol. 1: EasyChair, pp. 301-291, doi: https://doi.org/10.29007/4nk3.

[17] Y. Pan and L. Zhang, "Roles of artificial intelligence in construction engineering and management: A critical review and future trends," *Automation in Construction,* vol. 122, no. December 2020, pp. 103517-103517, 2021, doi: https://doi.org/10.1016/j.autcon.2020.103517.

[18] J. K. Pinto, *Project management : achieving competitive advantage*, Fourth edition. ed. Boston: Pearson (in English), 2016.

[19] S. Abbaspour, A. Martinez, G. S. Sandhu, and M. Nik-Bakht, "Automated Identification and Impact Quantification of Financial Budget Items from Construction Data," *Proceedings of the Canadian Society for Civil Engineering Annual Conference 2023, Volume 3,* pp. 1-15, 2024, doi: https://doi.org/10.1007/978-3-031-62170-3_1.

[20] Procore Technologies. "Survey Reveals how Technology Adoption Positively Impacts Performance." Procore Technologies. https://www.procore.com/press/78-of-corporate-and-public-real-estate-owners-over-budget-on-construction-projects-idc-survey (accessed September 19, 2025.

[21] M. Company. "The construction productivity imperative." https://www.mckinsey.com/capabilities/operations/our-insights/the-construction-productivity-imperative (accessed March 08, 2026.

[22] S. Dixit, S. N. Mandal, J. V. Thanikal, and K. Saurabh, "Evolution of studies in construction productivity: A systematic literature review (2006–2017)," *Ain Shams Engineering Journal,* vol. 10, no. 3, pp. 555-564, 2019, doi: 10.1016/j.asej.2018.10.010.

[23] A. Rivera, N. Le, J. Kashiwagi, and D. Kashiwagi, "Identifying the Global Performance of the Construction Industry," *Journal for the Advancement of Performance Information and Value,* vol. 8, no. 2, pp. 1-9, 2016, doi: https://doi.org/10.37265/japiv.v8i2.61.

[24] B. Golzarpoor, C. T. Haas, and D. Rayside, "Improving process conformance with Industry Foundation Processes (IFP)," *Advanced Engineering Informatics,* vol. 30, no. 2, pp. 143-156, 2016 2016, doi: https://doi.org/10.1016/j.aei.2016.02.005.

[25] Y. Zhu and G. Augenbroe, "A conceptual model for supporting the integration of inter-organizational information processes of AEC projects," *Automation in Construction,* vol. 15, no. 2, pp. 200-211, 2006, doi: https://doi.org/10.1016/j.autcon.2005.05.003.

[26] A. J. Martinez Lagunas and M. Nik-Bakht, "Process Mining, Modeling, and Management in Construction: A Critical Review of Three Decades of Research Coupled with a Current Industry Perspective," *Journal of Construction Engineering and Management,* vol. 150, no. 11, 2024, doi: https://doi.org/10.1061/jcemd4.coeng-14727.

[27] J. J. Shi, D. E. Lee, and E. Kuruku, "Task-based modeling method for construction business process modeling and automation," *Automation in Construction,* vol. 17, no. 5, pp. 633-640, 2008, doi: 10.1016/j.autcon.2007.10.010.

[28] NRC Canada. "Construction Sector Digitalization and Productivity Challenge program - National Research Council Canada." https://nrc.canada.ca/en/research-development/research-collaboration/programs/construction-digitalization-productivity-challenge-program (accessed August 22 2023.

[29] H. M. Consulting, "Process Mining: A Database of Applications 2021," 2021, vol. 2022. Accessed: June 15, 2025. [Online]. Available: https://www.tf-pm.org/news/a-2021-edition-of-the-extensive-list-of-process-mining-applications-has-been-released-by-hspi-s-p-a

[30] Celonis. "The Process Mining Buyer's Guide." Celonis. https://www.celonis.com/ebook/process-mining-buyers-guide/ (accessed April 10, 2022).

[31] S. S. van Schaijk, B. B. de Vries, and I. L. L. van Berlo, "Building Information Model (BIM) based process mining," 2016. [Online]. Available: https://research.tue.nl/en/studentTheses/building-information-model-bim-based-process-mining.

[32] H. Liu, "Economic performance associated with digitalization in Canada over the past two decades Economic and Social Reports," Statistics Canada, 2021. Accessed: October 10, 2025. [Online]. Available: https://doi.org/10.25318/36280001202100200001-eng

[33] J. Mollins and T. Taskin, "Digitalization: Productivity," Bank of Canada, 1914-0568, December 2023 2023. [Online]. Available: https://doi.org/10.34989/sdp-2023-17

[34] A. Charles Atkins *et al.*, "Rekindling US productivity for a new era," McKinsey Global Institute, December 2023 2023. [Online]. Available: https://www.mckinsey.com/mgi/our-research/rekindling-us-productivity-for-a-new-era

[35] A. J. Martinez Lagunas and M. Nik-Bakht, "Enabling Process Mining in the Construction Industry: An Event Log Schema for Change Management Process," in *Proceedings of the Canadian Society for Civil Engineering Annual Conference 2023, Volume 3*, Cham, S. Desjardins, G. J. Poitras, and M. Nik-Bakht, Eds., 2024: Springer Nature Switzerland, pp. 103-117, doi: https://doi.org/10.1007/978-3-031-62170-3_8.

[36] S. Shabani Ardakani and M. Nik-Bakht, "Functional Evaluation of Change Order and Invoice Management Processes under Different Procurement Strategies: Social Network Analysis Approach," *Journal of Construction Engineering and Management,* vol. 147, no. 1, 2021/1// 2021, doi: 10.1061/(asce)co.1943-7862.0001974.

[37] S. P. S. Padala, J. U. Maheswari, and H. Hirani, "Identification and classification of change causes and effects in construction projects," *International Journal of Construction Management,* vol. 22, no. 14, pp. 2788-2807, 2022, doi: 10.1080/15623599.2020.1827186.

[38] T. Pourrostam and A. Ismail, "Causes and Effects of Delay in Iranian Construction Projects," *International Journal of Engineering and Technology,* vol. 4, no. 5, pp. 598-601, 2012, doi: 10.7763/ijet.2012.v4.441.

[39] S. Ahmed and I. Arocho, "Analysis of cost comparison and effects of change orders during construction: Study of a mass timber and a concrete building project," *Journal of Building Engineering,* vol. 33, 2021/1// 2021, doi: 10.1016/j.jobe.2020.101856.

[40] S. A. Assaf and S. Al-Hejji, "Causes of delay in large construction projects," *International Journal of Project Management,* vol. 24, no. 4, pp. 349-357, 2006/5// 2006, doi: 10.1016/j.ijproman.2005.11.010.

[41] A. Emamifar, "Modeling Productivity losses Due to Change Orders," Master of Applied Science in Civil Engineering Thesis (Masters), Building, Civil and Environmental Engineering, Concordia University, Montreal, Quebec, Canada, 2019. [Online]. Available: https://spectrum.library.concordia.ca/id/eprint/985096/

[42] A. M. Eldeep, M. A. M. Farag, and L. M. Abd El-hafez, "Using BIM as a lean management tool in construction processes – A case study: Using BIM as a lean management tool," *Ain Shams Engineering Journal,* vol. 13, no. 2, 2022/3// 2022, doi: 10.1016/j.asej.2021.07.009.

[43] W. Ibbs, "Update on Quantitative Analysis of Change and Loss of Productivity," *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction,* vol. 13, no. 1, 2021, doi: https://doi.org/10.1061/(ASCE)LA.1943-4170.0000447.

[44] Marsh, GuyCarpenter, and O. Economics, "Future of Construction A Global Forecast for Construction to 2030," October 14, 2024 2021. Accessed: October 24, 2024. [Online]. Available: https://www.oxfordeconomics.com/resource/future-of-construction/

[45] S. Karimidorabati, "A Model for Implementing an Automated Change Management Process for Construction Mega-projects," 2014.

[46] A. J. Martinez Lagunas, S. Abbaspour, M. Nik-Bakht, and M. Ouf, "Digitalization and Process Management in Construction: A Current Industry Perspective," CMAA, 2025. Accessed: 2025/08/17. [Online]. Available: https://www.cmaanet.org/sites/default/files/resource/Digitalization%20FINAL.pdf

[47] R. Ndihokubwayo and T. Haupt, "VARIATION ORDERS ON CONSTRUCTION PROJECTS: VALUE ADDING OR WASTE?," *International Journal of Construction Project Management, IJCPM,* vol. 1, no. 2, 2009. [Online]. Available: https://www.researchgate.net/publication/45380450_Variation_Orders_on_Construction_Projects_Value-adding_or_Waste.

[48] V. T. Nguyen and S. T. Do, "Assessing the relationship chain among causes of variation orders, project performance, and stakeholder performance in construction projects," *International Journal of Construction Management,* vol. 23, no. 9, pp. 1592-1602, 2023, doi: 10.1080/15623599.2021.1988197.

[49] T. Ōno, *Toyota production system : beyond large-scale production*, [New York]: [Productivity Press], 1988.

[50] M. Fricke, "The knowledge pyramid: The dikw hierarchy," *Knowledge Organization,* vol. 46, pp. 33-46, 2019, doi: https://doi.org/10.5771/0943-7444-2019-1-33.

[51] W. M. P. Van der Aalst, M. Stoffele, and J. W. F. Wamelink, "Case handling in construction," *Automation in Construction,* vol. 12, no. 3, pp. 303-320, 2003, doi: https://doi.org/10.1016/S0926-5805(02)00106-1.

[52] M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management*. Berlin, Germany: Springer, 2018, p. 456.

[53] B. Golzarpoor, "Industry Foundation Processes (IFP): Theoretical and Practical Foundations for the Construction Industry," 2017. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/11172.

[54] S. A. Kelkar, *BUSINESS PROCESS MANAGEMENT: A Concise Study*. PHI Learning Pvt. Ltd. (in en), 2021, p. 418.

[55] ABPMP, *BPM CBOK*, 1st Editio ed. ABPMP International, 2013, p. 444.

[56] B. Dave, "Business process management - A construction case study," *Construction Innovation,* vol. 17, no. 1, pp. 50-67, 2017, doi: 10.1108/CI-10-2015-0055.

[57] G. r. Cachon and C. Terwiesch, *Matching supply with demand : an introduction to operations management*, Fourth edition. ed. (The McGraw-Hill/Irwin series in operations and decision sciences). New York, NY: McGraw-Hill Education (in English), 2019.

[58] C. Igwe, A. Hammad, and F. Nasiri, "Influence of lean construction wastes on the transformation-flow-value process of construction," *International Journal of Construction Management,* vol. 0, no. 0, pp. 1-7, 2020, doi: 10.1080/15623599.2020.1812153.

[59] M. Hammer, "The process audit," *Harvard Business Review,* vol. 85, no. 4, pp. 111-123, 2007.

[60] ISO, *Building information models: information delivery manual. Part 1 Methodology and format*, Second edition, 2016-05-01. ed. (International standard ; ISO 29481-1:2016(E)). Geneva, Switzerland: ISO (in English), 2016.

[61] J. Vom Brocke and J. Mendling, *Business Process Management Cases*. Springer Nature, 2018, p. 149.

[62] A. Anand, S. F. Wamba, and D. Gnanzou, "A literature review on business process management, business process reengineering, and business process innovation," *Lecture Notes in Business Information Processing,* vol. 153, pp. 1-23, 2013, doi: 10.1007/978-3-642-41638-5_1.

[63] C. N. Johnson, "The benefits fo PDCA," *Quality Progress,* vol. 35, no. 5, pp. 120-120, 2002 2002.

[64] M. L. R. Dumas, Marcello; Mendling, Jan; Reijers, Hajo A., *Business process management*. 2008, pp. 1-2.

[65] T. Frederick Winslow, *The Principles of Scientific Management*. Harper, 1919.

[66] M. Hammer and J. Champy, "Reengineering the corporation: A manifesto for business revolution," *Business Horizons,* vol. 36, no. 5, pp. 90-91, 1993 1993, doi: 10.1016/S0007-6813(05)80064-3.

[67] M. Hammer, "Reengineering Work : Don't Automate , Obliterate," pp. 1-12, 1990.

[68] T. M. Myers, "Re-engineering Rationale," *Journal of Management in Engineering,* vol. 11, no. August, pp. 11-13, 1995 1995, doi: 10.1061/(ASCE)0742-597X(1995)11:4(11).

[69] R. Al-Aomar, "A lean construction framework with six sigma rating," *International Journal of Lean Six Sigma,* vol. 3, no. 4, pp. 299-314, 2012/1// 2012, doi: 10.1108/20401461211284761.

[70] J. De Mast and J. Lokkerbol, "An analysis of the Six Sigma DMAIC method from the perspective of problem solving," *International Journal of Production Economics,* vol. 139, no. 2, pp. 604-614, 2012, doi: 10.1016/j.ijpe.2012.05.035.

[71] W. M. P. Van der Aalst, "On the Pareto Principle in Process Mining, Task Mining, and Robotic Process Automation.," in *DATA 2020 - Proceedings of the 9th International Conference on Data Science, Technology and Applications*, 2020: SciTePress, pp. 5-12, doi: 10.5220/0009979200050012.

[72] B. Ruecker and L. Strauch, *Enterprise Process Orchestration*. Hoboken, NJ: John Wiley & Sons, 2025, p. 304.

[73] W. Sharmak, I. R. J. Scherer, and I. U. Rüppel, *Dynamic Network planning in construction projects using configurable reference process models*. 2011, p. 9.

[74] M. A. Abdelmegid, V. A. González, M. O'Sullivan, C. G. Walker, M. Poshdar, and F. Ying, "The roles of conceptual modelling in improving construction simulation studies: A comprehensive review," *Advanced Engineering Informatics,* vol. 46, no. June, 2020, doi: 10.1016/j.aei.2020.101175.

[75] L. Costa, M. B. A. Barbosa, R. D. L. Baldam, and T. D. P. Coelho, "Challenges of Process Modeling in Architecture and Engineering to Execute Projects and Public Works," *Journal of Construction Engineering and Management,* vol. 145, no. 1, pp. 1-17, 2019, doi: https://doi.org/10.1061/(ASCE)CO.1943-7862.0001575.

[76] W. M. P. Van der Aalst, J. Carmona, J. Mylopoulos, S. Ram, M. Rosemann, and C. Szyperski, *Process Mining Handbook*. Springer, 2022.

[77] J. Ryan and C. Heavey, "Process modeling for simulation," *Computers in Industry,* vol. 57, no. 5, pp. 437-450, 2006 2006, doi: 10.1016/j.compind.2006.02.002.

[78] OMG, "Business Process Model and Notation (BPMN)," 2011 2011, doi: 10.1016/s0020-1693(97)05933-1.

[79] ISO, "ISO / IEC 19510 Information technology — Object Management Group Business Process," *International Organization for Standardization,* vol. E, 2013 2013. [Online]. Available: https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html.

[80] E. Marengo, W. Nutt, and M. Perktold, "CoPModL: Construction Process Modeling Language and Satisfiability Checking," *Information Systems,* no. xxxx, pp. 101457-101457, 2019 2019, doi: 10.1016/j.is.2019.101457.

[81] V. Sanvido, "Conceptual Construction Process Model," *Journal of Construction Engineering and Management,* vol. 114, no. 2, pp. 294-310, 1988 1988. [Online]. Available: https://doi-org.lib-ezproxy.concordia.ca/10.1061/(ASCE)0733-9364(1988)114:2(294).

[82] J. Gingele, S. J. Childe, and M. E. Miles, "A modelling technique for re-engineering business processes controlled by ISO 9001," *Computers in Industry,* vol. 49, no. 3, pp. 235-251, 2002 2002, doi: 10.1016/S0166-3615(02)00113-6.

[83] A. Benevolenskiy, "Ontology-based configuration of construction processes using process patterns," Technische Universität Dresden, 2015. [Online]. Available: https://tud.qucosa.de/api/qucosa%3A28204/attachment/ATT-0/?L=1

[84] C. Zhang, A. Hammad, T. M. Zayed, G. Wainer, and H. Pang, "Cell-based representation and analysis of spatial resources in construction simulation," *Automation in Construction,* vol. 16, no. 4, pp. 436-448, 2007 2007, doi: 10.1016/j.autcon.2006.07.009.

[85] Y. Lin, *Semantic Annotation for Process Models*. 2008, p. 1.

[86] M. Noguera, M. V. Hurtado, M. L. Rodríguez, L. Chung, and J. L. Garrido, "Ontology-driven analysis of UML-based collaborative processes using OWL-DL and CPN," *Science of Computer Programming,* vol. 75, no. 8, pp. 726-760, 2010, doi: 10.1016/j.scico.2009.05.002.

[87] M. De Leoni and W. M. P. Van Der Aalst, "Data-aware process mining: Discovering decisions in processes using alignments," *Proceedings of the ACM Symposium on Applied Computing,* no. January, pp. 1454-1461, 2013 2013, doi: 10.1145/2480362.2480633.

[88] L. Reinkemeyer, *Process Mining in Action*. 2020.

[89] S. Kouhestani and M. Nik-Bakht, "IFC-based process mining for design authoring," (in en), *Automation in Construction,* vol. 112, pp. 1-15, 2020, doi: 10.1016/j.autcon.2019.103069.

[90] Y. Pan and L. Zhang, "Automated process discovery from event logs in BIM construction projects," *Automation in Construction,* vol. 127, no. March, 2021, doi: https://doi.org/10.1016/j.autcon.2021.103713.

[91] W. M. P. Van der Aalst, "Process Mining: A 360 Degree Overview," in *Process Mining Handbook*, vol. 448: Springer Nature, 2022, pp. 3-34.

[92] W. M. P. Van der Aalst, *Process mining : discovery, conformance and enhancement of business processes*, Berlin: Springer, 2011. [Online]. Available: https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=372196.
[93] J. r. J. Becker, M. Kugeler, and M. Rosemann, *Process management : a guide for the design of business processes*. Berlin, New York: Springer (in English), 2003.
[94] S. M. Abourizk, S. A. Hague, and R. Ekyalimpa, *Construction simulation*. 2016, p. 378.
[95] Fluxicon. "Disco." Fluxicon. https://www.fluxicon.com/disco/ (accessed 2025/08/01.
[96] ProM. "ProM." http://www.promtools.org/doku.php?id=prom69 (accessed.
[97] Celonis. "Process Mining and Execution Management Software." https://www.celonis.com/ (accessed 2025-09-01.
[98] FIT. "PM4py - Process Mining for Python." Fraunhofer Institute for Applied Information Technology. https://pm4py.fit.fraunhofer.de/ (accessed September 14, 2023.
[99] bupaR. "Business Process Analysis." https://www.bupar.net/index.html (accessed.
[100] PMTk. "Process Intelligence Solutions (P.I.S.)." Process Intelligence Solutions (P.I.S.). https://processintelligence.solutions/pmtk/ (accessed 2025/08/01.
[101] Microsoft. "Process Mining Tool | Microsoft Power Automate." https://powerautomate.microsoft.com/en-us/process-mining/ (accessed September 14, 2023.
[102] Microsoft. "Overview of process mining and task mining in Power Automate." Microsoft. https://learn.microsoft.com/en-us/power-automate/process-advisor-overview (accessed 2025/08/02.
[103] IBM. "Fundamentals of IBM Process Mining." https://ibm-learning-skills-dev.github.io/badges/IBM_Process_Mining.html (accessed February 11, 2023).
[104] Palantir Technologies. "Foundry Process Mining & Automation." https://www.palantir.com/platforms/foundry/process-mining/ (accessed 2026/02/15.
[105] Apromore. "Full Spectrum Process Intelligence Platform | Apromore." https://apromore.com (accessed 2025-09-01.
[106] Gartner, "Magic Quadrant for Process Mining Tools," 2023. [Online]. Available: https://www.gartner.com/document/4192799?
[107] C. d. S. Garcia *et al.*, "Process mining techniques and applications – A systematic mapping study," *Expert Systems with Applications,* vol. 133, pp. 260-295, 2019, doi: 10.1016/j.eswa.2019.05.003.
[108] *Standard for eXtensible Event Stream (XES) for Achieving Interoperability in Event Logs and Event Streams*, 9781504424219, IEEE, 2016. [Online]. Available: https://ieeexplore.ieee.org/servlet/opac?punumber=7740856
[109] *OCEL: A Standard for Object-Centric Event Logs*, 9783030850814, A. F. Ghahfarokhi, G. Park, A. Berti, and W. M. P. van der Aalst, 2021. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-85082-1_16
[110] D. Fahland *et al.*, "Towards a Simple and Extensible Standard for Object-Centric Event Data (OCED) -- Core Model, Design Space, and Lessons Learned," 2024/10// 2024. [Online]. Available: http://arxiv.org/abs/2410.14495.
[111] Eindhoven University of Technology. "Analytics for Information Systems - AIS." https://www.win.tue.nl/ais/doku.php (accessed September 29, 2022).
[112] IEEE. "Institute of Electrical and Electronics Engineers." https://www.ieee.org/ (accessed September 28, 2022).

[113] IEEE, *IEEE Standard for eXtensible Event Stream for Achieving Interoperability in Event Logs and Event Streams*. 2023.

[114] J. C. A. M. Buijs, "Mapping Data Sources to XES in a Generic Way," in *Chelsea*, ed, 2010, pp. 123-123.

[115] C. Günther and E. Verbeek, "OpenXES Developer Guide," Eindhoven University of Technology, The Netherlands, 2012. [Online]. Available: https://www.xes-standard.org/_media/openxes/openxesdeveloperguide-1.8.pdf

[116] J. C. A. M. Buijs, "Mapping Data Sources to XES in a Generic Way," Master, Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, 2010. [Online]. Available: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Mapping+Data+Sources+to+XES+in+a+Generic+Way#0

[117] A. Berti *et al.*, "OCEL (Object-Centric Event Log) 2.0 Specification," 2023. [Online]. Available: https://www.ocel-standard.org/2.0/ocel20

[118] A. F. Ghahfarokhi, G. Park, A. Berti, and W. M. P. van der Aalst, *OCEL: A Standard for Object-Centric Event Logs* (Communications in Computer and Information Science). Springer International Publishing, 2021, pp. 169-175.

[119] A. Berti, M. Montali, and W. M. P. van der Aalst, "Advancements and Challenges in Object-Centric Process Mining: A Systematic Literature Review," 2023/11// 2023. [Online]. Available: http://arxiv.org/abs/2311.08795.

[120] D. Fahland, "Process Mining over Multiple Behavioral Dimensions with Event Knowledge Graphs," in *Lecture Notes in Business Information Processing*, vol. 448: Springer Science and Business Media Deutschland GmbH, 2022, pp. 274-319.

[121] C. W. Günther and W. M. P. van der Aalst, "Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics," in *Business Process Management*, Berlin, Heidelberg, G. Alonso, P. Dadam, and M. Rosemann, Eds., 2007, vol. 4714, no. Lecture Notes in Computer Science: Springer Berlin Heidelberg, in LNCS, pp. 328-343, doi: https://doi.org/10.1007/978-3-540-75183-0_24.

[122] S. J. J. Leemans, D. Fahland, and W. M. P. van der Aalst, "Discovering Block-Structured Process Models from Incomplete Event Logs," in *Application and Theory of Petri Nets and Concurrency.PETRI NETS 2014.* , Cham, G. Ciardo and E. Kindler, Eds., 2014, vol. 8489, no. Lecture Notes in Computer Science: Springer pp. 91-110, doi: https://doi.org/10.1007/978-3-319-07734-5_6.

[123] S. J. J. Leemans, D. Fahland, and W. M. P. van der Aalst, "Scalable Process Discovery with Guarantees," in *Enterprise, Business-Process and Information Systems Modeling. BPMDS EMMSAD 2015 2015.*, Cham, K. Gaaloul, R. Schmidt, S. Nurcan, S. Guerreiro, and Q. Ma, Eds., 2015, vol. 214, no. Lecture Notes in Business Information Processing: Springer, pp. 85-101, doi: https://doi.org/10.1007/978-3-319-19237-6_6.

[124] S. J. J. Leemans, E. Poppe, and M. T. Wynn, "Directly follows-based process mining: Exploration & a case study," in *International Conference on Process Mining, ICPM* 2019: Institute of Electrical and Electronics Engineers Inc., pp. 25-32, doi: https://doi.org/10.1109/ICPM.2019.00015.

[125] W. M. P. Van der Aalst, "On the Pareto Principle in Process Mining, Task Mining, and Robotic Process Automation," in *Proceedings of the 9th International Conference on Data Science, Technology and Applications* 2020: SCITEPRESS, in DATA, pp. 5-12, doi: https://doi.org/10.5220/0009979200050012.

[126] W. M. P. Van der Aalst, *Process Mining: Data Science in Action*, 2nd ed. Berlin, Heidelberg: Springer, 2016, pp. 1-477.

[127] F. Kog, R. J. Scherer, and A. Dikbas, "Verification and validation approach of BPMN represented construction processes," in *Creative Construction Conference 2014* 2014, pp. 265-271, doi: 10.1201/b12516-39. [Online]. Available: https://publications.cms.bgu.tum.de/proceedings/FBI2013/Kog_2013_FBI.pdf

[128] R. Cooper, G. Aouad, A. Lee, S. Wu, A. Fleming, and M. Kagioglou, *Process Management in Design and Construction*. Great Britain: Blackwell Publishin, 2005.

[129] H.-N. Su and P.-C. Lee, "Mapping knowledge structure by keyword co-occurrence: a first look at journal papers in Technology Foresight," *Scientometrics,* vol. 85, no. 1, pp. 65-79, 2010/10/01 2010, doi: 10.1007/s11192-010-0259-8.

[130] M. Hammer, "The process audit," *Harvard business review,* vol. 85, no. 4, pp. 111-9, 2007.

[131] S. Setti. "Project and process integration." Project Management Institute. https://www.pmi.org/learning/library/project-process-integration-combine-models-6801 (accessed Septembre 09, 2021).

[132] M. J. Horman and R. Kenley, "Quantifying Levels of Wasted Time in Construction with Meta-Analysis," *Journal of Construction Engineering and Management,* vol. 131, no. 1, pp. 52-61, 2005 2005, doi: 10.1061/(asce)0733-9364(2005)131:1(52).

[133] N. M. El-Gohary and T. E. El-Diraby, "Domain Ontology for Processes in Infrastructure and Construction," *Journal of Construction Engineering and Management,* vol. 136, no. 7, 2010 2010, doi: 10.1061/(asce)co.1943-7862.0000178.

[134] A. L. Stumpf, R. Ganeshan, S. Chin, and L. Y. Liu, "Object-Oriented Model for Integrating Construction Product and Process Information," *Journal of Computing in Civil Engineering,* vol. 10, no. 3, pp. 204-212, 1996, doi: 10.1061/(asce)0887-3801(1996)10:3(204).

[135] I. C. Wu, A. Borrmann, U. Beißert, M. König, and E. Rank, "Bridge construction schedule generation with pattern-based construction methods and constraint-based simulation," *Advanced Engineering Informatics,* vol. 24, no. 4, pp. 379-388, 2010 2010, doi: 10.1016/j.aei.2010.07.002.

[136] M. Nik-Bakht and T. E. El-Diraby, "Synthesis of Decision-Making Research in Construction," *Journal of Construction Engineering and Management,* vol. 141, no. 9, 2015 2015, doi: 10.1061/(asce)co.1943-7862.0000984.

[137] M. Nik-Bakht, "Analyzing Social Network Discussions of Infrastructure Projects-Towards a Bottom-Up Decision Making Environment," Doctor of Philosophy, Department of Civil Engineering, University of Toronto, 2015. [Online]. Available: https://www.proquest.com/docview/1999873329?pq-origsite=gscholar&fromopenview=true

[138] K. Sigalov and M. König, "Recognition of process patterns for BIM-based construction schedules," *Advanced Engineering Informatics,* vol. 33, pp. 456-472, 2017 2017, doi: 10.1016/j.aei.2016.12.003.

[139] S. Kouhestani, "Integration of BIM and Process Mining for Design Authoring Processes," no. March, pp. 39915587-39915587, 2019.

[140] C. Le, H. D. Jeong, and I. Damnjanovic, "Network Theory–Driven Construction Logic Knowledge Network: Process Modeling and Application in Highway Projects," *Journal of Construction Engineering and Management,* vol. 147, no. 10, 2021/10// 2021, doi: 10.1061/(asce)co.1943-7862.0002143.

[141] H. Liu, M. Al-Hussein, and M. Lu, "BIM-based integrated approach for detailed construction scheduling under resource constraints," *Automation in Construction,* vol. 53, pp. 29-43, 2015 2015, doi: 10.1016/j.autcon.2015.03.008.

[142] M. Kamleh, "Improving Hoist Performance during the Up-Peak of Tall Building Construction," 2014 2014. [Online]. Available: https://tspace.library.utoronto.ca/handle/1807/70181.

[143] A. Hosny, M. Nik-Bakht, and O. Moselhi, "Workspace planning in construction: non-deterministic factors," *Automation in Construction,* vol. 116, 2020/8// 2020, doi: 10.1016/j.autcon.2020.103222.

[144] S. Demirkesen and B. Ozorhon, "Impact of integration management on construction project management performance," *International Journal of Project Management,* vol. 35, no. 8, pp. 1639-1654, 2017 2017, doi: 10.1016/j.ijproman.2017.09.008.

[145] T. A. El-Diraby, C. Lima, and B. Feis, "Domain Taxonomy for Construction Concepts: Toward a Formal Ontology for Construction Knowledge," *Journal of Computing in Civil Engineering,* vol. 19, no. 4, pp. 394-406, 2005 2005, doi: 10.1061/(asce)0887-3801(2005)19:4(394).

[146] J. E. Kelley and M. R. Walker, "Critical-Path Planning and Scheduling," presented at the Eastern joint IRE-AIEE-ACM computer conference, 1959. [Online]. Available: https://doi.org/10.1145/1460299.1460318.

[147] M. Lu, W.-H. Chan, J.-P. Zhang, and M. Cao, "Generic Process Mapping and Simulation Methodology for Integrating Site Layout and Operations Planning in Construction," *Journal of Computing in Civil Engineering,* vol. 21, no. 6, pp. 453-462, 2007 2007, doi: 10.1061/(asce)0887-3801(2007)21:6(453).

[148] P. Tzortzopoulos, M. Sexton, and R. Cooper, "Process models implementation in the construction industry: A literature synthesis," *Engineering, Construction and Architectural Management,* vol. 12, no. 5, pp. 470-486, 2005, doi: 10.1108/09699980510627162.

[149] D. W. Halpin, "CYCLONE—Method for modeling job site processes," *Journal of the Construction Division,* vol. 103, no. 3, pp. 489-499, 1977 1977, doi: https://doi-org.lib-ezproxy.concordia.ca/10.1061/JCCEAZ.0000712.

[150] S. Moradi, F. Nasirzadeh, and F. Golkhoo, "A hybrid SD-DES simulation approach to model construction projects," *Construction Innovation,* vol. 15, no. 1, pp. 66-83, 2015, doi: 10.1108/CI-10-2013-0045.

[151] M. Raoufi and A. R. Fayek, "Fuzzy Monte Carlo Agent-Based Simulation of Construction Crew Performance," *Journal of Construction Engineering and Management,* vol. 146, no. 5, pp. 1-13, 2020, doi: 10.1061/(ASCE)CO.1943-7862.0001826.

[152] L. Ben-Alon and R. Sacks, "Simulating the behavior of trade crews in construction using agents and building information modeling," *Automation in Construction,* vol. 74, pp. 12-27, 2017, doi: 10.1016/j.autcon.2016.11.002.

[153] W. M. P. Van Der Aalst, "Process mining and simulation: A match made in heaven!," *Simulation Series,* vol. 50, no. 10, pp. 39-50, 2018, doi: 10.22360/summersim.2018.scsc.005.

[154] S. Ahn, P. S. Dunston, A. Kandil, and J. C. Martinez, "Process Mining Technique for Automated Simulation Model Generation Using Activity Log Data," *Computing in Civil Engineering,* pp. 636-643, 2015 2015, doi: 10.1061/9780784479247.079.

[155] R. Akhavian and A. H. Behzadan, "Knowledge-Based Simulation Modeling of Construction Fleet Operations Using Multimodal-Process Data Mining," *Journal of*

Construction Engineering and Management, vol. 139, no. 11, pp. 04013021-04013021, 2013, doi: 10.1061/(asce)co.1943-7862.0000775.

[156] R. Akhavian and A. H. Behzadan, "Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers," *Advanced Engineering Informatics,* vol. 29, no. 4, pp. 867-877, 2015, doi: 10.1016/j.aei.2015.03.001.

[157] R. Akhavian and A. H. Behzadan, "Coupling human activity recognition and wearable sensors for data-driven construction simulation," *Journal of Information Technology in Construction,* vol. 23, no. December 2016, pp. 1-15, 2018.

[158] M. A. Abdelmegid, V. A. González, M. O'Sullivan, C. G. Walker, M. Poshdar, and F. Ying, "Barriers to adopting simulation modelling in construction industry," *Automation in Construction,* vol. 111, no. December 2019, pp. 103046-103046, 2020 2020, doi: 10.1016/j.autcon.2019.103046.

[159] A. Ismail, Y. Srewil, and R. J. Scherer, "Integrated and collaborative process-based simulation framework for construction project planning," *International Journal of Simulation and Process Modelling,* vol. 12, no. 1, pp. 42-53, 2017 2017, doi: 10.1504/IJSPM.2017.082789.

[160] A. Benevolenskiy, K. Roos, P. Katranuschkov, and R. J. Scherer, "Construction processes configuration using process patterns," *Advanced Engineering Informatics,* vol. 26, no. 4, pp. 727-736, 2012, doi: https://doi.org/10.1016/j.aei.2012.04.003.

[161] C. N. Johnson. (2016) The Benefits of PDCA. *Quality Progress*. 45. Available: https://www.proquest.com/magazines/benefits-pdca/docview/1762043658/se-2

[162] M. S. M. Al Ahbabi, "Process protocol for the implementation of integrated project delivery in the UAE: A client perspective," in *University of Salford* vol. 55, ed, 2014, pp. 3909-3909.

[163] L. Zhang, Y. Pan, X. Wu, and M. J. Skibniewski, "Process Mining," in *Lecture Notes in Civil Engineering Artificial Intelligence in Construction Engineering and Management*: Springer Nature Singapore, 2021, pp. 147-171.

[164] M.-Y. Cheng and M.-H. Tsai, "Reengineering of Construction Management Process," *Journal of Construction Engineering and Management,* vol. 129, no. 1, pp. 105-114, 2003, doi: https://doi.org/10.1061/(ASCE)0733-9364(2003)129:1(105).

[165] M.-Y. Cheng, L.-C. Lien, C.-P. Huang, and C.-M. Huang, "Matrix Organization Process Reengineering for Construction Firms," *Journal of Management in Engineering,* vol. 31, no. 6, pp. 04014093-04014093, 2015 2015, doi: 10.1061/(asce)me.1943-5479.0000335.

[166] APQC, "Process Classification Framework," (in en), *APQC,* 2021 2021. [Online]. Available: https://www.apqc.org/resource-library/resource-listing/apqc-process-classification-framework-pcf-utilities-pdf-version.

[167] O. A. Ayodele, A. Chang-Richards, and V. González, "Factors Affecting Workforce Turnover in the Construction Sector: A Systematic Review," *Journal of Construction Engineering and Management,* vol. 146, no. 2, pp. 03119010-03119010, 2020 2020, doi: 10.1061/(asce)co.1943-7862.0001725.

[168] ISO. "ISO 16739-1:2018 Industry Foundation Classes (IFC)." ISO. https://www.iso.org/standard/70303.html (accessed.

[169] B. Golzarpoor, C. T. Haas, D. Rayside, S. Kang, and M. Weston, "Improving construction industry process interoperability with Industry Foundation Processes (IFP)," *Advanced*

Engineering Informatics, vol. 38, pp. 555-568, 2018, doi: https://doi.org/10.1016/j.aei.2018.09.001.

[170] N. M. El-Gohary and T. E. El-Diraby, "Dynamic Knowledge-Based Process Integration Portal for Collaborative Construction," *Journal of Construction Engineering and Management,* vol. 136, no. 3, pp. 316-328, 2010 2010, doi: 10.1061/(asce)co.1943-7862.0000147.

[171] E. Marengo, W. Nutt, and M. Perktold, "Construction Process Modelling," 2021 2021. [Online]. Available: http://bp-construction.inf.unibz.it:8080/ConstructionProcessModelling-beta/#/models.

[172] R. Akhavian and A. H. Behzadan, "Smartphone-based construction workers' activity recognition and classification," *Automation in Construction,* vol. 71, no. Part 2, pp. 198-209, 2016, doi: 10.1016/j.autcon.2016.08.015.

[173] B. Sherafat *et al.*, "Automated Methods for Activity Recognition of Construction Workers and Equipment: State-of-the-Art Review," *Journal of Construction Engineering and Management,* vol. 146, no. 6, pp. 03120002-03120002, 2020, doi: 10.1061/(asce)co.1943-7862.0001843.

[174] T. Slaton, C. Hernandez, and R. Akhavian, "Construction activity recognition with convolutional recurrent networks," *Automation in Construction,* vol. 113, no. August 2019, pp. 103138-103138, 2020, doi: 10.1016/j.autcon.2020.103138.

[175] M. Khazen, M. Nik-Bakht, and O. Moselhi, "Monitoring workers on indoor construction sites using data fusion of real-time worker's location, body orientation, and productivity state," *Automation in Construction,* vol. 160, 2024/4// 2024, doi: 10.1016/j.autcon.2024.105327.

[176] F. R. Correa, "Cyber-physical systems for construction industry," *Proceedings - 2018 IEEE Industrial Cyber-Physical Systems, ICPS 2018,* pp. 392-397, 2018 2018, doi: 10.1109/ICPHYS.2018.8387690.

[177] F. Amer and M. Golparvar-Fard, "Modeling dynamic construction work template from existing scheduling records via sequential machine learning," *Advanced Engineering Informatics,* vol. 47, no. October 2020, pp. 101198-101198, 2021 2021, doi: 10.1016/j.aei.2020.101198.

[178] S. V. Schaijk, "Building Information Model (BIM) based process mining: enabling knowledge reassurance and fact-based problem discovery within the Architecture, Engineering, Construction and Facility Management Industry," Eindhoven University of Technology, 2016. [Online]. Available: http://rgdoi.net/10.13140/RG.2.1.2276.9047

[179] K. M. Rashid and J. Louis, "Process Discovery and Conformance Checking in Modular Construction Using RFID and Process Mining," *ASCE Construction Research Congress 2020,* pp. 809-818, 2020, doi: https://doi-org.lib-ezproxy.concordia.ca/10.1061/9780784482865.068.

[180] A. F. Ghahfarokhi, A. Berti, and W. M. P. van der Aalst, "Object-Centric Process Mining Using Process Cubes," *International Journal of Mechanical and Industrial Engineering,* vol. 14, no. 11, pp. 478-485, 2020 2020. [Online]. Available: https://gitlab.com/Anahita-Farhang/process-cube%0Ahttps://publications.waset.org/10011600/object-centric-process-mining-using-process-cubes.

[181] A. J. Matinez Lagunas and M. Nik-Bakht, "Enabling Process Mining in the Construction Industry: An Event Log Schema for Change Management Process," in *CSCE/CRC*

International Construction Specialty Conference. (Manuscript accepted for publication), 2023, pp. 1-13.

[182] L. Holm, *Cost Accounting and Financial Management for Construction Project Managers*, 1st ed. Routledge, 2018, pp. 301-301.

[183] R. C. De Azevedo, R. T. De Oliveira Lacerda, L. Ensslin, A. E. Jungles, and S. R. Ensslin, "Performance Measurement to Aid Decision Making in the Budgeting Process for Apartment-Building Construction: Case Study Using MCDA-C," *Journal of Construction Engineering and Management,* vol. 139, no. 2, pp. 225-235, 2013/2// 2013, doi: 10.1061/(asce)co.1943-7862.0000587.

[184] T. Kropp, A. Bombeck, and K. Lennerts, "An Approach to Data Driven Process Discovery in the Cost Estimation Process of a Construction Company," presented at the 38th International Symposium on Automation and Robotics in Construction, 2021.

[185] J. H. Lee and J. S. Yi, "Predicting project's uncertainty risk in the bidding process by integrating unstructured text data and structured numerical data using text mining," *Applied Sciences (Switzerland),* vol. 7, no. 11, 2017/11// 2017, doi: 10.3390/app7111141.

[186] F. B. Muehlhausen and Y. Chen, *Practical construction accounting and financial management*. West Lafayette, Indiana: Purdue University Press (in English), 2023.

[187] S. Kouhestani and M. Nik-Bakht, "IFC-based process mining for design authoring," *Automation in Construction,* vol. 112, no. June 2019, 2020, doi: 10.1016/j.autcon.2019.103069.

[188] ASCE, "Construction Contract Claims, Changes, and Dispute Resolution," Reston, Virginia, USA, 2016. [Online]. Available: https://ascelibrary.org/doi/book/10.1061/9780784414293

[189] Government of Canada. "Standard on Embodied Carbon in Construction." Government of Canada. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32742&section=html (accessed January 30, 2024).

[190] L. Zhou, A. Rahman, T. Liu, and X. Zhang, "Blockchain Applications for Construction Contract Management: A Systematic Literature Review," *Journal of Construction Engineering and Management*, vol. 149, no. 1, doi: 10.1061/(ASCE)CO.1943-7862.0002428.

[191] A. J. Martinez Lagunas, S. Abbaspour, and M. Nik-Bakht, "NLP for Automated Discovery and Assessment of Dominant Construction and Maintenance Work Order Activities in MEP Projects," presented at the Construction Research Congress 2024: Advanced Technologies, Automation, and Computer Applications in Construction, Des Moines, Iowa, 2024. [Online]. Available: https://ascelibrary.org/doi/10.1061/9780784485262.018.

[192] BNQ, *Guide sur la gestion des contrats de construction dans un environnement ISO 9000*, 1re éd., 2006-03-17. ed. Sainte-Foy: Bureau de normalisation du Québec (in French), 2006.

[193] W. Gao, S. Lu, X. Zhang, Q. He, W. Huang, and B. Lin, "Impact of 3D modeling behavior patterns on the creativity of sustainable building design through process mining," *Automation in Construction,* vol. 150, 2023/6// 2023, doi: 10.1016/j.autcon.2023.104804.

[194] Qualtrics XM. "How to Carry Out Great Interviews in Qualitative Research." https://www.qualtrics.com/experience-management/research/qualitative-research-interview/ (accessed September 14, 2023.

[195] H. Luo, J. Liu, C. Li, K. Chen, and M. Zhang, "Ultra-rapid delivery of specialty field hospitals to combat COVID-19: Lessons learned from the Leishenshan Hospital project in

Wuhan," *Automation in Construction,* vol. 119, no. June, pp. 103345-103345, 2020, doi: 10.1016/j.autcon.2020.103345.

[196] F. L. Bennett, *The management of construction : a project life cycle approach*. Oxford: Butterworth-Heinemann (in English), 2003.

[197] S. Kang, "Semi-Automated Methods for Measuring Practice Conformance for Capital Projects," 2020 2020. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/16100.

[198] Q. Hao, W. Shen, J. Neelamkavil, and R. Thomas, "Change management in construction projects," in *Proceedings of the International Conference on Information Technology in Construction CIB W78*, 2008/07 2008: NRC, pp. 15-17. [Online]. Available: https://www.academia.edu/download/75905365/Change_management_in_construction_projec20211207-6156-131sn2h.pdf. [Online]. Available: https://www.academia.edu/download/75905365/Change_management_in_construction_projec20211207-6156-131sn2h.pdf

[199] Procore Technologies. "What is Procore? | Construction Management Software." https://www.procore.com/en-ca/what-is-procore (accessed 2025-09-01.

[200] ORACLE. "ORACLE NetSuite." Oracle. https://www.netsuite.com/portal/products/erp.shtml (accessed 2025-09-01.

[201] P. Technologies. "Resource Management Software | Procore." @procoretech. https://www.procore.com/resource-management (accessed 12/07/2025.

[202] Rhumbix. "RMBX." https://www.rhumbix.com/ (accessed September 26, 2022).

[203] Procore Technologies. "Change Orders - Workflow Diagrams." https://support.procore.com/products/online/user-guide/project-level/change-orders/workflow (accessed 2025-09-01.

[204] M. S. M. Al Ahbabi, "Process protocol for the implementation of integrated project delivery in the UAE: A client perspective," School of the Built Environment, University of Salford, 2014. [Online]. Available: https://salford-repository.worktribe.com/output/1426133

[205] G. Eken, G. Bilgin, I. Dikmen, and M. T. Birgonul, "A lessons-learned tool for organizational learning in construction," *Automation in Construction,* vol. 110, no. April 2019, pp. 102977-102977, 2020, doi: https://doi.org/10.1016/j.autcon.2019.102977.

[206] A. J. Martinez Lagunas and M. Nik-Bakht, "Enabling Process Mining in the Construction Industry: An Event Log Schema for Change Management Process," presented at the Canadian Society for Civil Engineering Annual Conference 2023, Moncton, Canada, 2024. [Online]. Available: https://link.springer.com/10.1007/978-3-031-62170-3_8.

[207] A. J. Martinez Lagunas and M. Nik-Bakht, "Analysis of XES and OCEL Data Schemas: Towards Multidimensional Process Mining of Intertwined Construction Processes," in *Proceedings of the 41st International Symposium on Automation and Robotics in Construction*, Lile, France, V. Gonzalez-Moret, J. Zhang, B. García de Soto, and I. Brilakis, Eds., 2024, vol. Construction Management Techniques, no. 2413-5844: International Association for Automation and Robotics in Construction (IAARC), pp. 444-451, doi: https://doi.org/10.22260/ISARC2024/0058.

[208] V. M. F. Pestana Ana Catarina, C. L. Alves Thaís da, and R. Barbosa André, "Application of Lean Construction Concepts to Manage the Submittal Process in AEC Projects," *Journal of Management in Engineering,* vol. 30, no. 4, p. 05014006, 2014, doi: https://doi.org/10.1061/(ASCE)ME.1943-5479.0000215.

[209] L. Reinkemeyer, S. Cham, Ed. *Process Intelligence in Action: Taking Process Mining to the Next Level*. 2024, pp. 1-251.

[210] M. L. Van Eck, X. Lu, S. J. J. Leemans, and W. M. P. Van der Aalst, "PM2: A Process Mining Project Methodology," *Advanced Information Systems Engineering. CAiSE 2015. Lecture Notes in Computer Science* vol. 9097, pp. 520-521, 2015, doi: https://doi.org/10.1007/978-3-319-19069-3_19.

[211] Oracle. "Aconex Construction Management Platform." https://www.oracle.com/ca-en/construction-engineering/aconex/ (accessed 2025-09-01.

[212] Autodesk. "Construction Management Software | Autodesk Construction Cloud." Autodesk. https://construction.autodesk.com/ (accessed 2025-09-01.

[213] SAP. "What is SAP? | Definition and Meaning." @SAP. https://www.sap.com/about/company/what-is-sap.html (accessed 2025-09-01.

[214] CMiC. "CMiC: Construction Management Software | CMiC." https://cmicglobal.com/ (accessed April 07, 2025).

[215] Sage. "Construction Accounting Software." Sage. https://www.sage.com/en-ca/sage-construction/ (accessed 2025-09-01.

[216] Intuit Quickbooks. "QuickBooks Canada® Official Site | Smart Tools. Better Business." https://quickbooks.intuit.com/ca/ (accessed 2025-09-01

[217] "Cloud Computing Services | Microsoft Azure." https://azure.microsoft.com/en-us (accessed May 22, 2025).

[218] Microsoft. "SQL Server Management Studio (SSMS)." https://learn.microsoft.com/en-us/ssms/sql-server-management-studio-ssms (accessed May 22, 2025).

[219] Y. Wang, S. Liao, Z. Gong, F. Deng, and S. Yin, "Enhancing Construction Management Digital Twins Through Process Mining of Progress Logs," *Sustainability (Switzerland),* vol. 16, no. 22, 2024, doi: https://doi.org/10.3390/su162210064.

[220] L. Van Den Ingh, R. Eshuis, and S. Gelper, "Assessing performance of mined business process variants," *Enterprise Information Systems,* vol. 15, no. 5, pp. 676-693, 2021, doi: https://doi.org/10.1080/17517575.2020.1746405.

[221] P. T. G. Hornix, "Performance Analysis of Business Processes through Process Mining," Eindhoven University of Technology, 2007. [Online]. Available: https://research.tue.nl/files/47012708/625917-1.pdf

[222] M. Abduh and F. Adam, "PERFORMANCE OF REQUEST FOR INFORMATION (RFI) MANAGEMENT," 2016. [Online]. Available: https://web.archive.org/web/20200709103329id_/https://www.isec-society.org/ISEC_PRESS/ASEA_SEC_03/pdf/C-6_v1_43.pdf.

[223] M. Rahman and S. H. Han, "Linear Scheduling Method-Based Multi-Objective Optimization for Off-Site Construction Manufacturing," *Automation in Construction,* vol. 167, 2024, doi: https://doi.org/10.1016/j.autcon.2024.105712.

[224] V. V. Denisov, E. Belkina, and D. Fahland, "BPIC'2018: Mining Concept Drift in Performance Spectra of Processes," 2018, doi: https://doi.org/10.4121/uuid:3301445f-95e8-4ff0-98a4-901f1f204972.

[225] A. Laghouag, F. b. Zafrah, M. R. N. M. Qureshi, and A. A. Sahli, "Eliminating Non-Value-Added Activities and Optimizing Manufacturing Processes Using Process Mining: A Stock of Challenges for Family SMEs," *Sustainability (Switzerland),* vol. 16, no. 4, 2024, doi: https://doi.org/10.3390/su16041694.

[226] M. Gunduz and A. F. Naser, "Cost based Value Stream Mapping as a sustainable construction tool for undergraound pipeline construction projects," *Sustainability (Switzerland),* vol. 9, no. 12, 2017, doi: https://doi.org/10.3390/su9122184.

[227] N. Hughes and L. Ap, "Impact and Control of RFIs on Construction projects," 2013. Accessed: 2025-09-01. [Online]. Available: https://www.cmaanet.org/sites/default/files/resource/Impact%20%26%20Control%20of%20RFIs%20on%20Construction%20Projects.pdf

[228] S. Mohamed, P. A. Tilley, and S. N. Tucker, "QUANTIFYING THE TIME AND COST ASSOCIATED WITH THE REQUEST FOR INFORMATION (RFI) PROCESS IN CONSTRUCTION." Accessed: 2025-09-01. [Online]. Available: https://research.usq.edu.au/item/q22q6/quantifying-the-time-and-cost-associated-with-the-request-for-information-rfi-process-in-construction

[229] M. Nik Bakht, "Analyzing Social Network Discussions of Infrastructure Projects—Towards a Bottom-Up Decision Making Environment," Ph.D., University of Toronto (Canada), Canada -- Ontario, CA, 10186355, 2015. [Online]. Available: https://search.proquest.com/openview/5665d79058d521d240eb1b9927a4f55c/1?pq-origsite=gscholar&cbl=18750

[230] L. Zhang and B. Ashuri, "BIM log mining: Discovering social networks," (in en), *Automation in Construction,* vol. 91, pp. 31-43, 2018/07// 2018, doi: https://doi.org/10.1016/j.autcon.2018.03.009.

[231] W. M. P. Van der Aalst, H. A. Reijers, and M. Song, "Discovering social networks from event logs," *Computer Supported Cooperative Work,* vol. 14, no. 6, pp. 549-593, 2005, doi: 10.1007/s10606-005-9005-9.

[232] W. Van Woensel and S. Motie, "NLP4PBM: a systematic review on process extraction using natural language processing with rule-based, machine and deep learning methods," *Enterprise Information Systems,* 2024, doi: https://doi.org/10.1080/17517575.2024.2417404.

[233] E. González López de Murillas, H. A. Reijers, and W. M. P. van der Aalst, "Connecting databases with process mining: a meta model and toolset," *Software and Systems Modeling,* vol. 18, no. 2, pp. 1209-1247, 2019, doi: https://doi.org/10.1007/s10270-018-0664-7.

[234] S. Da Silva and C. Boton, "Digital Integration in Construction: A Case Study on Common Data Environment Implementation for a Metro Line Project," *Infrastructures*, vol. 10, no. 10*,* p. 266doi: 10.3390/infrastructures10100266.

[235] P. Technologies. "Procore Developers." https://developers.procore.com/ (accessed May 22, 2025).

[236] *IEEE Standard for eXtensible Event Stream (XES) for Achieving Interoperability in Event Logs and Event Streams*, IEEE, 2023. [Online]. Available: https://standards.ieee.org/ieee/1849/10907/

[237] J. C. A. M. Buijs, B. F. Van Dongen, and W. M. P. Van der Aalst, "Quality dimensions in process discovery: The importance of fitness, precision, generalization and simplicity," *International Journal of Cooperative Information Systems,* vol. 23, no. 1, 2014, doi: https://doi.org/10.1142/S0218843014400012.

[238] J. N. Adams, S. J. van Zelst, T. Rose, and W. M. P. van der Aalst, "Explainable concept drift in process mining," *Information Systems,* pp. 102177-102177, 2023, doi: https://doi.org/10.1016/j.is.2023.102177.

[239] A. International, *Total Cost Management (TCM) Framework: An Integrated Approach to Project, Program, and Portfolio Management*, 2nd Edition ed. AACE International, 2015.

[240] M. Msawil, D. Greenwood, and M. Kassem, "A Systematic evaluation of blockchain-enabled contract administration in construction projects," *Automation in Construction,* vol. 143, 2022, doi: 10.1016/j.autcon.2022.104553.

[241] Procore Technologies. "Rest API Overview - Procore Developers." Procore Technologies, Inc. https://developers.procore.com/reference/rest/docs/rest-api-overview (accessed May 23, 2025).

[242] M. Afzal, J. K. W. Wong, and A. Ahmadian Fard Fini, "Towards digital approach for managing request for information (RFI) in construction projects: a literature review," *Construction Innovation,* 2024, doi: https://doi.org/10.1108/CI-05-2023-0113.

[243] A. S. Hanna, E. J. Tadt, and G. C. Whited, "Request for Information: Benchmarks and Metrics for Major Highway Projects," *Journal of Construction Engineering and Management,* vol. 138, no. 12, pp. 1347-1352, 2012, doi: https://doi.org/10.1061/(ASCE)CO.1943-7862.0000554.

[244] M. Das, X. Tao, Y. Liu, and J. C. P. Cheng, "A blockchain-based integrated document management framework for construction applications," *Automation in Construction,* vol. 133, p. 104001, 2022, doi: https://doi.org/10.1016/j.autcon.2021.104001.

[245] A. S. Erri Pradeep, T. W. Yiu, Y. Zou, and R. Amor, "Blockchain-aided information exchange records for design liability control and improved security," *Automation in Construction,* vol. 126, p. 103667, 2021, doi: https://doi.org/10.1016/j.autcon.2021.103667.

[246] S. Mohamed, P. A. Tilley, and S. N. Tucker, "Quantifying the time and cost associated with the request for information (RFI) process in construction," *Construction Innovation: information, process, management,* vol. 7, no. 1, pp. 35-50, 1999. [Online]. Available: https://research.usq.edu.au/item/q22q6/quantifying-the-time-and-cost-associated-with-the-request-for-information-rfi-process-in-construction.

[247] D. Papajohn and M. El Asmar, "Impact of Alternative Delivery on the Response Time of Requests for Information for Highway Projects," *Journal of Management in Engineering,* vol. 37, no. 1, 2021, doi: https://doi.org/10.1061/(ASCE)ME.1943-5479.0000866.

[248] J. J. Kim, A. L. Petrov, J. Lim, and S. Kim, "Comparing Cost Performance of Project Delivery Methods Using Quantifiable RFIs: Cases in California Heavy Civil Construction Projects," *International Journal of Civil Engineering,* vol. 20, no. 3, pp. 323-335, 2022, doi: https://doi.org/10.1007/s40999-021-00658-0.

[249] J. G. Zack, Jr., "Requests for information--Use, abuse, and control," presented at the Cost Engineering, Morgantown, 1999. [Online]. Available: https://www.proquest.com/docview/220430146?pq-origsite=gscholar&fromopenview=true&sourcetype=Scholarly%20Journals.

[250] I. S. M. Sandoval, J. P. A. O. Bernardo, S. R. Dionisio, and A. H. Aquino, "Application of Information and Communication Technologies (ICT) to optimize the request for information (RFI) process in construction projects," in *2023 Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI)*, 2023, pp. 1-6, doi: https://doi.org/10.1109/CONIITI61170.2023.10324094.

[251] A. A. Aibinu, C. Simon, F. Valerie, and P. V. and Serra, "Necessary evils: Controlling Requests for Information (RFIs) to reduce cost and improve margins," *Construction*

Research and Innovation, vol. 9, no. 4, pp. 103-108, 2018, doi: https://doi.org/10.1080/20450249.2018.1546976.

[252] A. J. Martinez Lagunas, S. Abbaspour, and M. Nik-Bakht, "NLP for Automated Discovery and Assessment of Dominant Construction and Maintenance Work Order Activities in MEP Projects," *Construction Research Congress 2024,* pp. 168-178, 2024, doi: https://doi.org/10.1061/9780784485262.018.

[253] S. Kouhestani and M. Nik-Bakht, "IFC-based process mining for design authoring," *Automation in Construction,* vol. 112, no. June 2019, pp. 1-15, 2020, doi: https://doi.org/10.1016/j.autcon.2019.103069.

[254] M. Afzal, J. K. Wai Wong, and A. A. Fard Fini, "Unlocking Insights: Analysing Construction Issues in Request for Information (RFI) Documents with Text Mining and Visualisation," in *IEEE International Conference on Automation Science and Engineering*, 2023: IEEE Computer Society, doi: https://doi.org/10.1109/CASE56687.2023.10260517.

[255] Y. Neziha and E. Esin, "Enhancing RFI Analysis in Construction Projects: A Comparative Study of Text Clustering Methods and Visualization Techniques," presented at the Proceedings of the 2024 European Conference on Computing in Construction, 2024.

[256] A. S. Bhat, "Data visualization of requests for information to support construction decision-making," Master of Applied Science, Civil Engineering, 2017. [Online]. Available: https://open.library.ubc.ca/collections/24/items/1.0344010

[257] Ö. Ceyhun and E. Esin, "Metadata Extraction of RFIs Using Natural Language Processing and Machine Learning Algorithms," *Proceedings of the 2024 European Conference on Computing in Construction,* vol. 5, 2024, doi: https://doi.org/10.35490/EC3.2024.287.

[258] R. Panahi, J.-P. Kivlin, and J. Louis, "Request for Information (RFI) Recommender System for Pre-Construction Design Review Application Using Natural Language Processing, Chat-GPT, and Computer Vision," *Computing in Civil Engineering 2023,* pp. 159-166, 2024, doi: https://doi.org/10.1061/9780784485224.020.

[259] W. Mao, Y. Zhu, and I. Ahmad, "Applying metadata models to unstructured content of construction documents: A view-based approach," *Automation in Construction,* vol. 16, no. 2, pp. 242-252, 2007, doi: https://doi.org/10.1016/j.autcon.2006.05.005.

[260] S. Adamtey, "Cost and time performance analysis of progressive design-build projects," (in English), *Journal of Engineering, Design and Technology,* vol. 19, no. 3, pp. 686-697, 2021, doi: https://doi.org/10.1108/JEDT-05-2020-0164.

[261] A. A. Aibinu, S. Carter, V. Francis, and P. Vaz-Serra, "Request for information frequency and their turnaround time in construction projects: A data-analytic study," *Built Environment Project and Asset Management,* vol. 10, no. 1, pp. 1-15, 2020, doi: https://doi.org/10.1108/BEPAM-10-2018-0130.

[262] W. Ibrahim Michael, S. Hanna Awad, S. Russell Jeffrey, S. Abotaleb Ibrahim, and H. El-adaway Islam, "Quantitative Analysis of the Impacts of Out-of-Sequence Work on Project Performance," *Journal of Construction Engineering and Management,* vol. 146, no. 8, p. 04020090, 2020, doi: https://doi.org/10.1061/(ASCE)CO.1943-7862.0001876.

[263] Y. Wang, S. Liao, Z. Gong, F. Deng, and S. Yin, "Enhancing Construction Management Digital Twins Through Process Mining of Progress Logs," *Sustainability (Switzerland),* vol. 16, no. 22, 2024, doi: https://doi.org/10.3390/su162210064.

[264] AECOM. "Elite Squad Puts the Tech into Giant Construction Projects - Engineering.com." Engineering. https://www.engineering.com/elite-squad-puts-the-tech-into-giant-construction-projects/ (accessed 2025/07/19.

[265] W. M. P. Van der Aalst, "A practitioner's guide to process mining: Limitations of the directly-follows graph," in *Procedia Computer Science*, 2019, vol. 164: Elsevier B.V., pp. 321-328, doi: https://doi.org/10.1016/j.procs.2019.12.189.

[266] Mindzie. "Mindzie - AI-Driven Process Mining and Intelligence." Mindzie. https://mindzie.com/ (accessed 2025/08/24.

[267] N. Yilmaz, E. Ergen, and A. M. Citipitioglu, "Enhancing RFI management in construction through machine learning-driven predictive models," *Smart and Sustainable Built Environment,* pp. 1-21, 2025, doi: 10.1108/SASBE-05-2025-0254.

[268] L. Zimmermann, F. Zerbato, and B. Weber, "What makes life for process mining analysts difficult? A reflection of challenges," *Software and Systems Modeling,* vol. 23, no. 6, pp. 1345-1373, 2024, doi: 10.1007/s10270-023-01134-0.

[269] W. M. P. van der Aalst, "No AI Without PI ! Object-Centric Process Mining as the Enabler for Generative, Predictive, and Prescriptive Artificial Intelligence," *Intelligent and Fuzzy Systems. INFUS 2025.,* pp. 28-38, 2025, doi: https://doi.org/10.1007/978-3-031-97985-9_4.

[270] W. C. Wang, S. W. Weng, S. H. Wang, and C. Y. Chen, "Integrating building information models with construction process simulations for project scheduling support," *Automation in Construction,* vol. 37, pp. 68-80, 2014 2014, doi: 10.1016/j.autcon.2013.10.009.

[271] J. P. Pantouvakis and A. Panas, "Computer simulation and analysis framework for floating caisson construction operations," *Automation in Construction,* vol. 36, pp. 196-207, 2013 2013, doi: 10.1016/j.autcon.2013.04.003.

[272] Z. S. Moussavi, A. Akbarnezhad, and D. Rey, "Optimization of concrete placing operation based on competing carbon footprint, cost and production rate objectives," *Engineering, Construction and Architectural Management,* vol. 25, no. 7, pp. 938-957, 2018 2018, doi: 10.1108/ECAM-05-2017-0080.

[273] H.-x. Liu *et al.*, "Simulation of Mobile Falsework Utilization Methods in Bridge Construction," presented at the Proceedings of the 2012 Winter Simulation Conference (WSC), 2012. [Online]. Available: https://doi.org/10.1109/WSC.2012.6465325

[274] L. Wu, W. Ji, and S. M. AbouRizk, "Bayesian Inference with Markov Chain Monte Carlo–Based Numerical Approach for Input Model Updating," *Journal of Computing in Civil Engineering,* vol. 34, no. 1, pp. 04019043-04019043, 2020 2020, doi: 10.1061/(asce)cp.1943-5487.0000862.

[275] T. Zayed, M. R. Sharifi, S. Baciu, and M. Amer, "Slip-Form Application to Concrete Structures," *Journal of Construction Engineering and Management,* vol. 134, no. 3, pp. 157-168, 2008, doi: 10.1061/(asce)0733-9364(2008)134:3(157).

[276] M. Arashpour, R. Wakefield, N. Blismas, and E. W. M. Lee, "Analysis of disruptions caused by construction field rework on productivity in residential projects," *Journal of Construction Engineering and Management,* vol. 140, no. 2, pp. 1-12, 2014, doi: 10.1061/(ASCE)CO.1943-7862.0000804.

[277] A. Alvanchi, R. Azimi, S. Lee, S. M. Abourizk, and P. Zubick, "Off-site construction planning using discrete event simulation," *Journal of Architectural Engineering,* vol. 18, no. 2, pp. 114-122, 2012 2012, doi: 10.1061/(ASCE)AE.1943-5568.0000055.

[278] K. P. Kisi, N. Mani, E. M. Rojas, and E. T. Foster, "Optimal productivity in labor-intensive construction operations: Pilot study," *Journal of Construction Engineering and Management,* vol. 143, no. 3, pp. 1-11, 2017, doi: 10.1061/(ASCE)CO.1943-7862.0001257.

[279] M. Lu, "Simplified discrete-event simulation approach for construction simulation," *Journal of Construction Engineering and Management,* vol. 129, no. 5, pp. 537-546, 2003, doi: 10.1061/(ASCE)0733-9364(2003)129:5(537).

[280] M. Mawlana, F. Vahdatikhaki, A. Doriani, and A. Hammad, "Integrating 4D modeling and discrete event simulation for phasing evaluation of elevated urban highway reconstruction projects," *Automation in Construction,* vol. 60, pp. 25-38, 2015 2015, doi: 10.1016/j.autcon.2015.09.005.

[281] J. Larsson, W. Lu, J. Krantz, and T. Olofsson, "Discrete Event Simulation Analysis of Product and Process Platforms: A Bridge Construction Case Study," *Journal of Construction Engineering and Management,* vol. 142, no. 4, pp. 04015097-04015097, 2016, doi: 10.1061/(asce)co.1943-7862.0001093.

[282] R. Akhavian and A. H. Behzadan, "An integrated data collection and analysis framework for remote monitoring and planning of construction operations," *Advanced Engineering Informatics,* vol. 26, no. 4, pp. 749-761, 2012, doi: 10.1016/j.aei.2012.04.004.

[283] K. M. Shawki, K. Kilani, and M. A. Gomaa, "Analysis of earth-moving systems using discrete-event simulation," *Alexandria Engineering Journal,* vol. 54, no. 3, pp. 533-540, 2015, doi: 10.1016/j.aej.2015.03.034.

[284] D. E. Lee, C. Y. Yi, T. K. Lim, and D. Arditi, "Integrated simulation system for construction operation and project scheduling," *Journal of Computing in Civil Engineering,* vol. 24, no. 6, pp. 557-569, 2010, doi: 10.1061/(ASCE)CP.1943-5487.0000061.

[285] S. Karimidorabati, "A Model for Implementing an Automated Change Management Process for Construction Mega-projects," Civil and Environmental Engineering, University of Waterloo, Waterloo, Ontario, 2014. [Online]. Available: https://uwspace.uwaterloo.ca/items/a6c92914-d212-47c7-888d-4947cb831b16

[286] S. Kang, "Semi-Automated Methods for Measuring Practice Conformance for Capital Projects," 2020. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/16100

[287] J. Niu and R. R. A. Issa, "Developing taxonomy for the domain ontology of construction contractual semantics: A case study on the AIA A201 document," *Advanced Engineering Informatics,* vol. 29, no. 3, pp. 472-482, 2015/8// 2015, doi: 10.1016/j.aei.2015.03.009.

[288] R. Akhavian and A. H. Behzadan, "Coupling human activity recognition and wearable sensors for data-driven construction simulation," *Journal of Information Technology in Construction,* vol. 23, no. December 2016, 2018 2018. [Online]. Available: https://itcon.org/papers/2018_01-ITcon-Akhavian.pdf.

APPENDICES

Table S1. Process Modeling in Discrete Event Simulation Studies - Construction Operations

Criteria	[270]	[271]	[272]	[142]	[273]	[274]
Simulation Type	4D + DES	DES	DES + Optimization	DES	DES	DES
Title	Integrating Building Information Models with Constructio n Process Simulations for Project Scheduling Support	Computer Simulation and Analysis Framework for Floating Caisson Construction Operations	Optimization of concrete placing operation based on competing carbon footprint, cost and production rate objectives	Improving Hoist Performance during Up-peak of Tall Building Construction	Simulation of mobile falsework utilization methods in bridge construction	Bayesian Inference with Markov Chain Monte Carlo-Based Numerical Approach for Input Model Updating
Construction Operations	Forming Concreting Hoisting Reinforcing	Slip forming Reinforcing Concreting	Concreting	Hoisting	Falsework	Earthmoving Operation
Targets (Decision Variables)	Project Total Duration	Project Total Duration Productivity Rates Total Unit Cost Resource Utilization	Project Cost Productivity Rates Carbon Emissions	Hoist's cycle time % of workers waiting Average waiting time Delivery of workers	Installation Cycle Times Resource Utilization Project Cost	Truck Cycle Time (Univariate)
Type of Data Collection Method	Manual	Manual & Semi-Automated	Manual	Manual	Manual & Semi-Automated	Manual
Data Source Type	Subjective	Recorded & Subjective	Collected Records	Collected Records & Subjective	Recorded & Subjective	Subjective
Data Source Details	Contractor's feedback	Direct Observations Time-lapse video Subject expert's consultation Project information	Direct Observations Previous Records & Studies	Historical Manufacturer's Data , Direct Observations and records operations and expert's feedback	Video and Site daily direct observation records Expert's consultation	Expert's knowledge and historical observations
Selected Model Simulation SW tool	Stroboscope	Stroboscope	SimEvents & Matlab	Simphony.NET	Simphony.NET	Simphony.NET
Does it involve user-written code?	Yes- Visual Basic	Yes- Visual Basic	Yes- Matlab	Yes- execute within simphony.net	Yes- execute within simphony.net	NA
Dependency Type among	Finish to Start	Finish to Finish Start to Start	Finish to Start	Finish to Start	Finish to Finish Start to Start	Finish to Start

Tasks/Activitie s						
Independence & homogeneity Statistical check	✗	✓ (Krustal-Wallis test)	✗	✗	✗	✗
Manual or Selected SW tool for distribution fitting of model inputs	Manual	"@Risk" ("Bestfit")	Matlab's "dfittool"	Manual	"@Risk" ("Bestfit")	Manual
Fitted Statistical Distributions- Input Model	Betta-PERT Distribution	Uniform Triangular Beta	Gamma Beta	Constant Inter-arrival/step/function. Empirical Distribution	"Durations- "@Risk" Cost- Normal	Beta for hauling duration Constant for other's durations
Goodness of fit	✗	✓ (Chi-square; K-S test; Anderson-Darling; Q-Q & P-P graphs)	✓ (Kolmogorov-Smirnov test)	✓ Cumulative Distribution Function	✓ ("@Risk checks for KS Test or Chi-Square")	✓ Visual CDF
Model Verification (Details)	✓ (logical outcomes vs. supervisor's expertise)	✓ (model's logic with subject experts)	✓ (model's results vs. optimization results consistency)	✓ (by tracing model components and their interactions vs. data from site & expert's opinion)	✓ (by tracing model elements and by observing statistics)	✓ (model's logic with subject experts)
Model Calibration (Details)	✗	✗	✗	✗	✗	✓
Model Validation (Details)	✗	✓ (model's performance results after 30 runs vs. actual field data)	✗	✓ (model's functioning vs. related previous studies and site data) % accuracy is not provided	✓ (model's performance results after 100 runs vs. actual field data)	✗
Sensitivity Analysis / Perturbation (Details)	✗	✓ (model's outcome performance through variation of critical model parameters)	✓ (model's outcome performance through variation of critical model parameters)	✓ (model's result's analysis through variation of model parameters)	✓ (model's result's analysis through variation of resource utilization)	✓ (100 runs variating the activities durations due to the learning effect of activities)

Table S1. Process Modeling in Discrete Event Simulation Studies - Construction Operations (Continuation)

Criteria	[275]	[276]	[277]	[278]	[279]	[280]
Simulation Type	DES	DES	DES	DES	SDESA	4D + DES
Title	Slip-Form Application to Concrete Structures	Analysis of Disruptions Caused by Construction Field Rework on Productivity in Residential Projects	Off-Site Construction Planing Using Discrete Event Simulation	Optimal Productivity in Labor-Intensive Construction Operations: Pilot Study	Simplified Discrete-Event Simulation Approach for Construction Simulation	Integrating 4D modeling and discrete event simulation for phasing evaluation of elevated urban highway reconstruction projects
Construction Operations	Slip-form, reinforcing and concreting	House building operation	Off-site structural steel fabrication	Electrical Lighting Replacement	Asphalt paving	Elevated Urban Highway projects: Construction and demolition of concrete box girder bridge
Targets (Decision Variables)	Productivity Slip-form Rate Resource Utilization	Productivity based on rework analysis	Project Total Duration Resource Utilization	Optimal Productivity	Arrival time of five asphalt trucks Cycle Time	Productivity Rates Costs per Conctractor
Type of Data Collection Method	Manual	Manual	Manual	Manual	Manual	Manual
Data Source Type	Subjective	Recorded	Subjective	Recorded and Subjective	Recorded	"Subjective (Contractors)
Data Source Details	Expert's knowledge of case study projects	Direct Observations from Site	Subject expert's consultation from workshops	"Direct Observations on Site	Expert's knowledge of case strudy projects	Direct Observations from Site
Selected Model Simulation SW tool	MicroCYCLONE	Arena 14.5	Simphony.NET	Arena	HKCONSIM	MicroCYCLONE
Does it involve user-written code?	✗	Yes- SIMAN	Yes- Execute Operator through Simphony.net	✗	Yes- Visual Basic	Yes- Visual Basic
Dependency Type among Tasks/Activities	Finish to Start	Finish to Start	Finish to Start Finish to Finish	Finish to Start	Finish to Start	Finish to Start

Independence & homogeneity Statistical check	✗	✗	✗	✓ (Paired-tmeans comparison test)	✗	✗
Manual or Selected SW tool for distribution fitting of model inputs	Manual	Input Analyzer Tool	Manual	Input Analyzer Tool	VIBES	Manual
Fitted Statistical Distributions-Input Model	Triangular	Exponential (Rework Data)	Constant Uniform Triangular	Exponential Weibull Gamma	Triangular	Triangular Normal
Goodness of fit	✗	✓ (test statistics and min square error)	✗	✓ (Chi-square; K-S test; p-value)	✗	✓ (Visual Cumulative Distribution Function)
Model Verification (Details)	✓ (model's logic with subject experts)	✓ (factorial ANOVA)	✓ (model's logic with subject experts)	✓ (model's sequence vs actual sequence in the field)	✓ (model's sequence vs actual sequence in the field-5 runs @ scenario)	✓ (mode'ls logic with subject experts/enginee rs)
Model Calibration (Details)	✗	✗	✗	✗	✗	✗
Model Validation (Details)	✓ (model's performance results vs actual field data)	✓ (DES Model Simulation vs Mathematical Modeling)	✓ (model's performance results vs actual field data)	✓ (Face validity technique (model's logic vs actual)	✓ (model's performance results vs actual field data)	✓ (model's performance results after 1000 runs)
Sensitivity Analysis / Perturbation (Details)	✓ (model's outcome performance through variation of critical input model parameters)	✓ (20,100+ runs were carried out)	✓ (model's outcome performance through variation of critical input model parameters)	✗	✗	✓ (1000 runs for validating the productivity rates and costs)

Table S1. Process Modeling in Discrete Event Simulation Studies - Construction Operations (Continuation)

Criteria	**[281]**	**[282]**	**[283]**	**[157]**	**[284]**
Simulation Type	DES	DES	DES	DES + Mobile Tracking	DES
Title	Discrete Event Simulation Analysis of Product and Process Platforms: A Bridge Construction Case Study	An integrated data collection and analysis framework for remote monitoring and planning of construction operations	Analysis of earth-moving systems using discrete-event simulation	Coupling Human Activity Recognition and Wearable Sensors for Data-Dirven Construction Simulation	Integrated Simulation System for Construction Operation and Project Scheduling
Construction Operations	Bridge Construction	Earthmoving Operation	Earthmoving Operation	Carpentry works	Earthmoving Operation
Targets (Decision Variables)	Total construction time Resource Utilization	Idle (waiting time) of construction equipment	Resource Utilization Idle Time Cycle Time Productivity	Cycle Time Idle Time	Cycle Time Resource Utilization Project Completion Time Cost
Type of Data Collection Method	Manual	Automated	Manual	Automated	Manual
Data Source Type	Recorded and Subjective	Recorded Real Time Data Collection	Subjective	Recorded	Subjective
Data Source Details	Interviews with project manager and subject experts Project Documentation Direct Site Observations Workshops	Accelerometer and Gyroscope Video Recordings	From previous study	Accelerometer and Gyroscope (Sensor Kinetics Pro App) Videotape	Fictitious Historical Data
Selected Model Simulation SW tool	SIMIO	STROBOSCOPE & VITASCOPE	Arena	STROBOSCOPE	SimEvent & Web-Cyclone
Does it involve user-written code?	✗	Yes- C++ .NET	✗	✗	Yes - MATLAB (MathWorks 2007a)

Dependency Type among Tasks/Activities	Finish to Start Finish to Finish Start to Start	Finish to Start	Finish to Start	Finish to Start	Finish to Start
Independence & homogeneity Statistical check	✗	✓ (trend of the collected data and outlier removal)	✗	✓ (t-test)	✓ (t-test)
Manual or Selected SW tool for distribution fitting of model inputs	Manual	Manual	Input Analyzer Tool	Input Analyzer Tool	Manual
Fitted Statistical Distributions-Input Model	Triangular	Normal	✗	Triangular (Sawing) Triangular (Loading) Gamma (Hauling) Beta (Unloading) Gama (Returning) Normal (Hammering) Beta (Turning the Wrench)	Triangular (activities' durations) Weibull (OPC) Normal (PPT)
Goodness of fit	✗	✗ Subjective (Expert Knowledge)	✗	✓ (Chi-square; K-S test; p-value-histograms)	✓ (Chi-square test-log likelihood test)
Model Verification (Details)	✓ (model’s logic with subject experts)	✓ (model’s logic based on 3D visualization /animation as pre-processing of data)	✓ (model’s logic vs 2D animation)	✓ (model’s sequence vs actual sequence in the experimental setting)	✗
Model Calibration (Details)	✗	✓ (Fine-tuning the model based on real time data)	✗	✓ (Machine Learning- ANN-SVM-KNN-Logistic Regression-DT)	✗
Model Validation (Details)	✓ (model's performance results after 1000 runs vs actual filed data)	✓ (model's performance results vs actual results from 3D orientation trackers in	✓ (model's performance results vs previous study results)	✓ (model's performance results (50 replications) vs observed values)	✓ (model's performance results vs previous study results)

		an indoor laboratory)			
Sensitivity Analysis / Perturbation (Details)	✓ (model's outcome performance through variation of critical input model parameters in the relational database)	✗	✗	✗	✓ (model's outcome performance through variation of critical input model parameters- 120 iterations)

Table S2. Process identification vs. process model type – administrative processes

Modeled Construction Process	“As-Planned” Process	“As-Happened” or “As-Is” Process	“To-Be” Process Forward-Looking
Request For Information		Golzarpoor, et al. [24]; Golzarpoor, et al. [169]	
Construction Change Order Management	Karimidorabati [285];	[286] Golzarpoor [53]	Karimidorabati [285]
Procurement/ Purchasing	Cheng and Tsai [164]; Cheng, et al. [165]; Shi, et al. [27]		
Project Delivery Management	Costa, et al. [75]		
Contract Management	[287]		
Cost Estimation		[184]	
Drawing/Design Approval	Schaijk [178]	Kouhestani and Nik-Bakht [89]	
Organizational Management	Sanvido [81]		

Table S3. Process identification vs. process model type – construction operational processes

Modeled Construction Process	“As-Planned” Process	As-Happened” or “As-Is” Process	“To-Be” Process Forward-Looking
Cast-in-place concrete elements	Benevolenskiy, et al. [160]; Sigalov and König [138]		
Precast concrete elements	Benevolenskiy [83]; Sigalov and König [138] [51]	[51]	

Electrical Installation	Sigalov and König [138]		
Sanitary Installation	Sigalov and König [138]		
Lay Floor	Marengo, et al. [80]		
Install Wooden Windows	Marengo, et al. [80]		
Assembly Columns	Benevolenskiy, et al. [160]	Correa [176]	
Masonry Walls		Correa [176]	
Heating, ventilation, and air conditioning installations		Van Der Aalst et al., (2003)	
Panelized building construction	Liu, et al. [141]		Liu, et al. [141]
Architectural & Structural Works	Schaijk [178]; Pan and Zhang [90]	Schaijk [178]; Pan and Zhang [90]	
Modular Construction		Rashid and Louis [179]	
Forming			Wang, et al. [270]
Hoisting			Kamleh [142];
Reinforcing			Pantouvakis and Panas [271]; Wang, et al. [270]
Falsework			Liu, et al. [273]
Earthmoving			Wu, et al. [274]
Maintenance and Error Handling		Schaijk [178];	
Slip forming			Pantouvakis and Panas [271]
Off-site Structural Steel Fabrication			[277]
Asphalt paving			[279]
Phasing of urban highway reconstruction			[280]
Bridge construction			Wu, et al. [135]
Carpentry works			[288]

Table S4. Comparison of ontology-based studies – ‘top-down’ process modeling approach

References (Ontology-based Modeling)	**Ontology Name**	**Axioms**	**Actors & Roles**	**Resources**	**Mechanisms**	**Processes**	**Products**	**Modalities/Clusters**	**Relationships**	**Systems**	**Attributes**	**Rules or Constraints**	**Events**	**Actions**	**Projects**	**Activity Locations**	**Families /Modality Views**	**Technical Topics**	**Model Level of Abstraction**
[145]	e-COGNOS (knowledge maNagement across prOjects and between enterpriSes)	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓			✓			✓	
[133] *[170]*	IC-PRO-Onto (Infrastructure and Construction PROcess Ontology)	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓		
[160] *[83]*	Ontology-based modeling	✓		✓	✓	✓	✓		✓		✓	✓				✓			
(Golzarpoor et al., 2016 *[53]* *[169]*	IFP (Industry Foundation Processes)	✓	✓			✓			✓		✓	✓		✓	✓				✓

Table S5. Aggregated analysis of ‘top-down’ PMM studies

Study	Underlaying Ontology?	Pattern-based Modeling?	Automated Model Validation	Process Analysis	Implemented WMS	Quantitative KPIs	Redesigned Model?	Implementation of redesigned model	Process Monitoring	Process Transparency	Conformance Checking	Process Interoperability	Process Reusability	Process Collaboration	Process Improvement	Process Flexibility/Configuration
Costa et al., 2019				✓	✓		✓	✓		✓			✓		✓	
Golzarpoor et al., 2016	✓	✓	✓	✓	✓					✓	✓		✓	✓		✓
Golzarpoor et al., 2018	✓	✓	✓	✓	✓							✓		✓		✓
Golzarpoor, 2017	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓		✓
Cheng & Tsai, 2003				✓		✓										
Cheng et al., 2015				✓	✓	✓	✓		✓				✓		✓	✓
Sanvido, 1988				✓		✓				✓	✓		✓	✓		
El-Gohary & El-Diraby, 2010b	✓		✓	✓						✓		✓	✓	✓		✓
El-Gohary & El-Diraby, 2010a	✓			✓	✓					✓		✓	✓	✓		✓
El-Diraby et al., 2005	✓				✓					✓		✓	✓	✓		✓
Benevolenskiy et al., 2012	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓		✓
Benevolenskiy, 2015	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓		✓

Table S5. Aggregated analysis of 'top-down' PMM studies (continuation)

Authors	Process Modeling Representation Language	Reported used software tools, instruments, or technologies	Addressed Major Problems	Modeled Process? As-Is Process Modeling	Reported Model's Validation Method
Costa et al., 2019	BPMN	Not Specified	Inefficient process management in public projects	Manage Project Delivery	Experts Feedback/ Focus Groups
Golzarpoor et al., 2016	BPMN & FOL	Skelta BPM & EPPM	Inefficient processes & time-consuming workflow customizations	Request For Information	Expert Feedback & Functional Demonstration
Golzarpoor et al., 2018	RDF, BPMN & .NET Framework	Microsoft Workflow Foundation & Windows Communication Foundation	Project delays & poor process non-proprietary interoperability	Change Requests Request For Information	Expert Feedback & Functional Demonstration
Golzarpoor, 2017	BPMN	EPPM	Non-conformant & non-interoperable processes	Change Requests Request For Information	Expert Feedback & Functional Demonstration
Cheng & Tsai, 2003	IDEF	Not Specified	Redundancy of business operations Poor Process Planning	Procurement/Purchasing	Customer & Expert's Feedback
Cheng et al., 2015	EPC	ARIS	High turnover rates of workers Inefficient Process Management	Procurement/Purchasing	KPIs-Process Value
Sanvido, 1988	IDEF	Not Specified	Management Ineffectiveness Failing Management Systems Low Productivity	Operational support to craftsman	CONNIE & Comparison between Reference Models and Actual Projects
El-Gohary & El-Diraby, 2010b	Tree-like Structure, NLP & FOL	Protégé-OWL	Poor Knowledge-based process conceptualization and coordination	Process Management	Competency Questions, Previous Ontologies; Expert Feedback & & RACER

El-Gohary & El-Diraby, 2010a	XML (Extensible Markup Language)	Protégé-OWL-Prototype SW Portal (Ontology-Merger)	The complexity of dynamic process integration	The urban transportation planning process	Experts Feedback/ Focus Group
El-Diraby et al., 2005	DAML (Agent Markup Language)	OWL (e-CKMI portal)	Lack of consistent semantic representation of construction knowledge	Knowledge Management Taxonomy	Competency Questions, Interviews, Workshop & Expert Feedback
Benevolenskiy et al., 2012	RDF (Resource Description Framework)	OWL-Process Configurator SW	Low quality and lack of automation in process modeling; Poor process patterns reusability	Produce Concrete Columns Assembly Columns	Functional Application (1- project case)
Benevolenskiy, 2015	RDF & BPMN	OWL; Drools, OptaPlanner & Process Configurator	Time-consuming, unstructured & rigid process models	Produce Precast Columns Produce Precast Slabs	Functional Application (2 - project cases)

Table S6. Aggregated analysis of ‘bottom-up’ PMM studies

Authors	**Modeled Process? As-Is Process Modeling**	**Relational Data (Δ) or Ontology-based (֍)?**	**Pattern-based Modeling?**	**Automated Model Validation**	**Process Analysis**	**Implemented WMS**	**Quantitative KPIs**	**To-Be Model Process Redesign?**	**Process Implementation**	**Process Monitoring**	**Process Transparency**	**Conformance Checking**	**Process Interoperability**	**Process Reusability**	**Allows for Dynamic Process Representation?**	**Process Collaboration**	**Process Re-engineering/ Improvement**	**Process Flexibility/ Configuration**
Shi et al., 2008	Procurement/ Purchasing	Δ		✓	✓						✓			✓	✓			✓
Sigalov & König, 2017	Cast-in-place columns Electrical and Sanitary Installation	Δ	✓	✓	✓		✓				✓	✓		✓				
Marengo et al., 2019	Lay Floor & Install Wooden Windows	Δ		✓	✓	✓	✓				✓	✓			✓	✓		✓
Correa, 2018	Assembly Columns and masonry walls	Δ			✓	✓					✓							
Amer & Golpavar-Fard, 2021	Various construction patterns	Δ	✓	✓	✓		✓				✓			✓	✓			
Liu et al., 2015	On-site construction of panelized building projects	Δ	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓

Table S6. Aggregated analysis of ‘bottom-up’ PMM studies (continuation)

Authors	Process Modeling Representation Language	Reported used software tools, instruments, or technologies	Addressed Major Problems	Reported Model's Evaluation/Validation Method
Shi et al., 2008	Traditional Workflow	Computer Coded Tasks (different levels of automation per task according to actors involved)	Not enough process automation at the task level & Poor Process Management	Not Specified
Sigalov & König, 2017	Feature Graphs Process Pattern Workflow	BIM-based scheduling software Branch and Bound Visualization	Poor identification and reusability of generic process patterns	Similarity Metric & Expert's Assessment
Marengo et al., 2019	CoPModL	CoPMod Process Modeling Platform	Need for a specific process modeling language for CEM with formal semantics	Satisfiability checking algorithm vs. NuSMV model checker
Correa, 2018	Petri nets	Ultra -wideband (UWB) system	Lack of integration between BIM and real project site information	Not Specified
Amer & Golpavar-Fard, 2021	Process Pattern Workflow	FastText & Projector TensorFlow	Manual creation of process templates based on previous projects is time-consuming and error-prone	Perturbation analysis through vanilla architectures
Liu et al., 2015	Process Pattern Workflow	Simphony.NET	Overlooked resource-constrains Poor integration of BIM with construction processes	Expert Feedback & Functional Demonstration

Table S7. Aggregated analysis of process mining studies

Study Details	Design		Construction			Operation
Author	[89]	[178]	[178]	[90]	[179]	[178]
Title of the Study	IFC-based process mining for design authoring	Case study: Construction design process mining	Building Information Model (BIM) based process mining	Automated process discovery from event logs in BIM construction projects	Process Discovery and Conformance Checking in Modular Construction Using RFID and Process Mining	Case study: Process mining with facility management data
Main Data Source	BIM Model in Revit	Systems Engineer IT system database	IFC ArchiCAD Design BIM Models Construction Planning Schedule IFC 4D Models in Synchro As-Built 3D Point Clouds -UAV	ArchiCAD Design BIM Models Construction Schedule IFC 4D Models in Synchro XES Event Logs	Radio Frequency Identification (RFID) Bldg. Component Tracking from Database	Planon Facility Management System
Developed Algorithms / Specific Contributions	IFC Archiving Dynamo Algorithm IFC Logging Dynamo Algorithm	No Developed Algorithms	Event Log Service Planning Consult Service	No Developed Algorithms	RFID and Process Mining Framework	No Developed Algorithms
Information Exchange File Extension Details	RVT→(DYN→IFC) →CSV→XES	IT System→CSV	PNL→(IFC+XML) → SP→ (IFC 4D VS PLY) → BIMSERVER→ CSV→XES	CSV→XES	Raw Data Putty Streaming→ CSV	Panon FM→CSV
Process Model Types	As-Planned (Design Authoring Process) As-Happened (Discovered Design Operational Model)	As-Is Process Model (Design Specification Process)	As-Planned & As-Built (Construction Operational Process)	As-Planned & As-Built (Construction Operational Process)	Predefined Production Plan vs. As-Is Process (Assembly Line Process)	Standard FM Process vs. As-Is Discovered Process (Maintenance Error Handling Process)
Applied Discovery Process Mining Algorithms	Inductive Miner	Fuzzy Miner	Fuzzy Miner	Inductive Miner Fuzzy Miner	Inductive Miner	Fuzzy Miner
Process Modeling Notations	Process Trees; Petri Nets; ProM Process Animation	Process Maps	Process Trees; Petri Nets; ProM Process Animation; Process map	Process Trees; Petri Nets; ProM Process Animation; Process map	Petri Nets/Process Tree	Process Maps

Applied Process Mining Techniques	Process Discovery Conformance Checking (Bottlenecks & Deviations) Social Network Analysis	Process Discovery Conformance Checking (Bottlenecks & Deviations) Social Network Analysis	Process Discovery Conformance Checking (Bottlenecks & Deviations) Social Network Analysis	Process Discovery Conformance Checking (Bottlenecks & Deviations) Social Network Analysis	Process Discovery Conformance Checking (Deviations & Bottlenecks Process Enhancement	Process Discovery Conformance Checking (Deviations & Bottlenecks)
Process Mining Perspectives	Control-Flow Perspective (Phase & Activity) Time Perspective Organizational Perspective (Actor-Centric)	Control-Flow Perspective (Activity Flow) Organizational Perspective (Actor-Centric)	Control-Flow Perspective (Activity-Level) Organizational Perspective (Actor-Centric)	Organizational Perspective (Actor-Centric) Time Perspective	Control-Flow Perspective (Component-Level) Time Perspective	Control-Flow Perspective (Steps for error handling) Time- Cost Perspective
Quality KPIs	SNA Metrics	SNA Metrics	SNA Metrics	Fitness & Precision & SNA Metrics	Not Mentioned	Not Mentioned
Functional Demonstration Details - Verification	2-story building (299elements)	Experimental Study Large Civil Project in the Netherlands	Fictive Project Home Building	Not Applicable	Laboratory Experiment Experimental Setup Of an assembly line	Experimental Visual Test in Disco
Validation	Real Case Study; Usability Test ; Customer Acceptance	Experiment's Results vs. Expert's Opinion	Exported Elements with Event Log Service; Real Case Study with assumed process actors Visual Validation of Planning Consult	Real Case Study with Assumed Process Actors	Laboratory Experiment Experimental Setup	Actual Process Mining Results vs. Standard Process and Expert's Opinion
Case Study Details	3-story hotel (7923 elements)	Civil Project	"Schependomlaan" 3-story residential building project	"Schependomlaan" 3-story residential building project	Modular Construction Facility	Facility Management of the MC Hospital in Rotterdam
Event Log Datapoints	Event Log I: 310 events Event Log II: 30,308 events	139,018 datapoints	3661 datapoints	3661 datapoints	13 datapoints	2414 datapoints
Event Log Attributes	4 Attributes	4 Attributes	12 attributes	12 attributes	4 Attributes	5 Attributes
Involved Modeling Disciplines	Architectural Structural MEP	Civil	Architectural and Structural	Architectural and Structural	Assembly Line 5 Units	Examples included: Sanitary, Electric & Plumbing

Supporting Tools/Platforms	Design: Revit IFC Export: Dynamo Process Mining: ProM	IT System Process Mining: Disco & My Invenio	Design: ArchiCAD As-Built: UAV & D4AR Tool As-Planned IFC Models: Synchro; BIM Server; Process Mining: Disco & My Invenio	Design: ArchiCAD As-Planned IFC Models: Synchro BIM Server Process Mining: ProM &Disco	RFID System Data Streaming: Putty Process Mining: ProM	FM Platform: Panon FM Process Mining: ProM &Disco
Main Target Group	BIM Managers	Process Engineers Project Managers Contractors	BIM and Process Mining Researchers Industry Practitioners	Project Managers	Process Managers	Facility Managers
Main Enhancement Targets	Designer Productivity	Process Efficiency Process Transparency	Process Planning	Process Efficiency	Productivity Process Improvement	Process Discovery Process Improvement

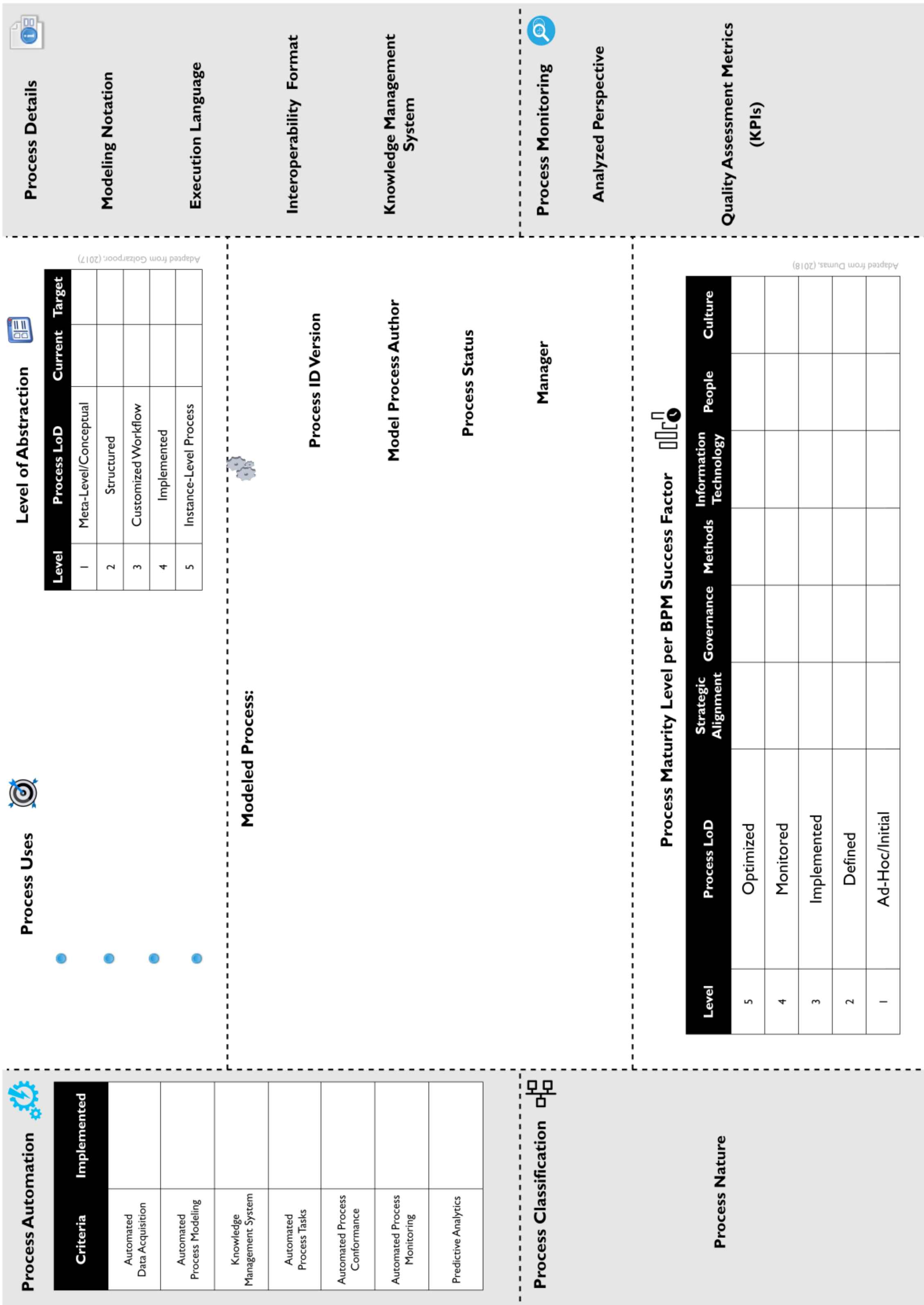


Figure S1. Business Process Model Canvas – Blank Template

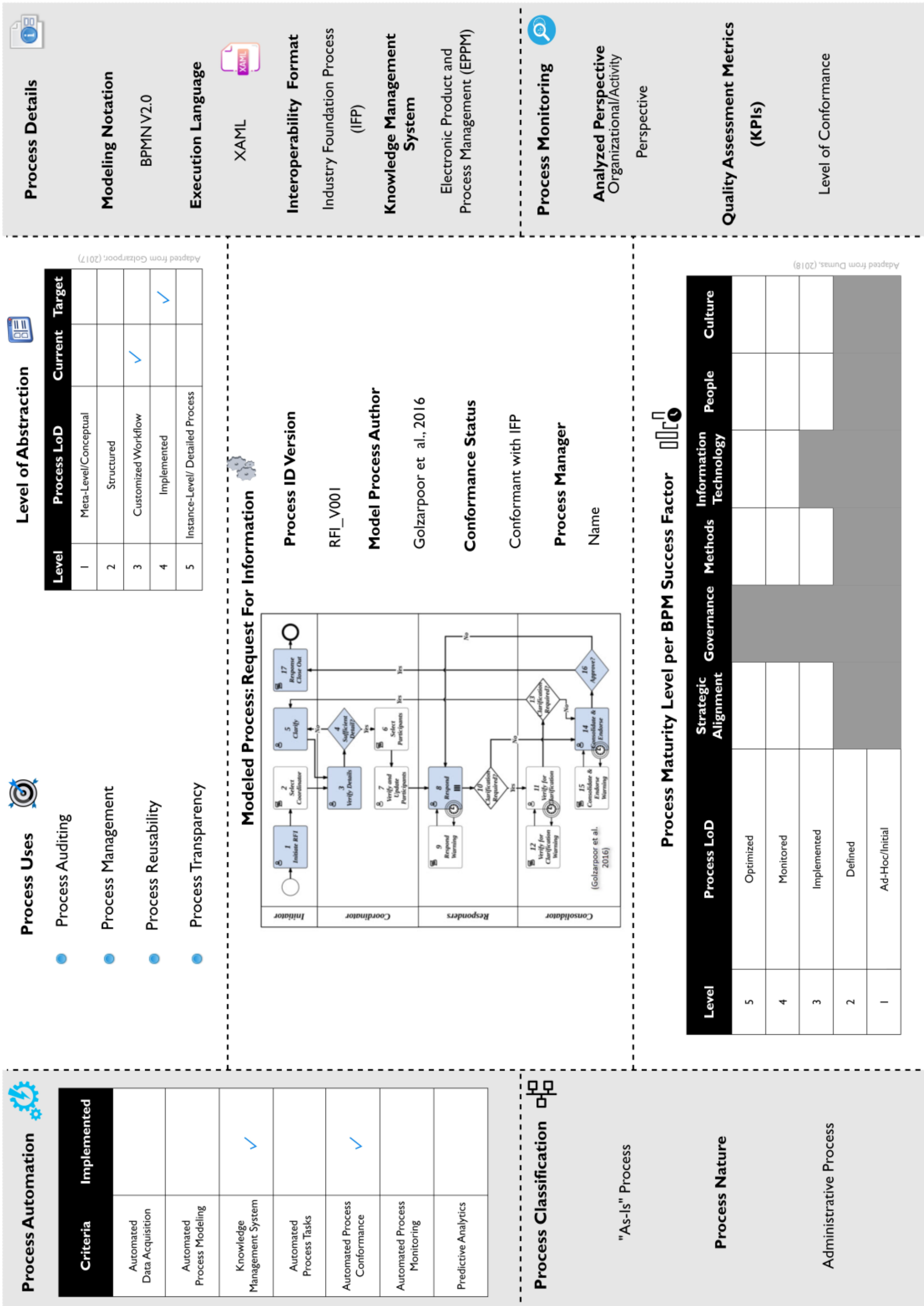


Figure S2. Administrative "As-Planned" process model in conformance with IFP.

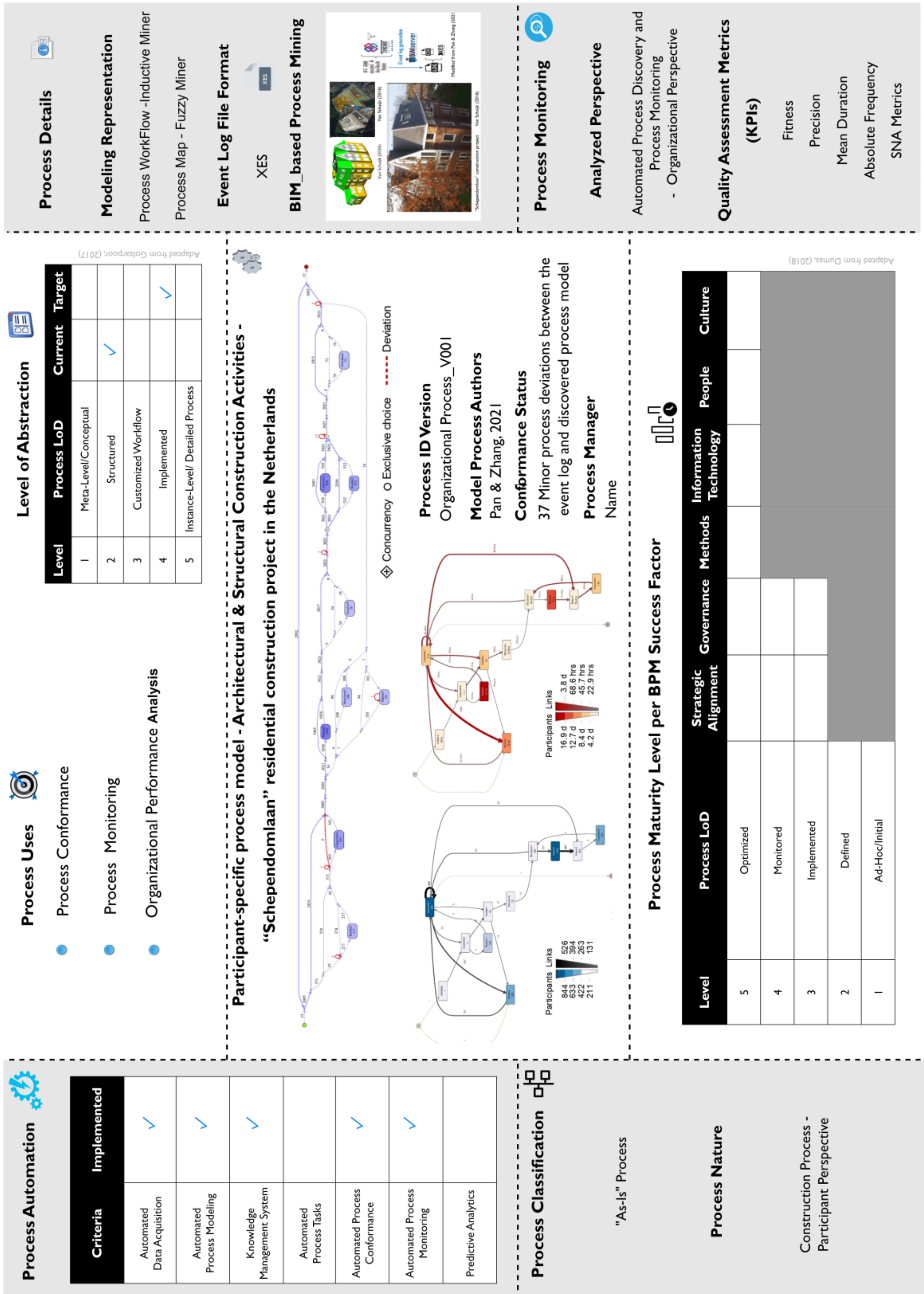


Figure S3. Construction operational "As-is/As-happened" process model.

Table S8. Perceived Significance of Construction Business Processes Based on Industry Expert Input

Phase	Process	Q1 Importance				Q2 Impact on project's completion timeline		
		L (%)	**M (%)**	**H (%)**	**N/E (%)**	**L (%)**	**M (%)**	**H (%)**
Tendering	Bidding			**(100%)**		(34%)		**(66%)**
Pre-Construction	Contract Management			**(83%)**	(17%)	(20%)	(40%)	**(40%)**
	Design Approval		(33%)	**(50%)**	(17%)	(20%)	(20%)	**(60%)**
	Submittal/ Transmittal Approval		(17%)	**(83%)**			(17%)	**(83%)**
	Procurement Management	(17%)	**(50%)**	(17%)	(17%)	(20%)	(40%)	**(40%)**
Construction	Labor/Resource Management	(17%)		**(66%)**	(17%)			
	Material Management		(34%)	**(66%)**		(17%)	(33%)	**(50%)**
	Construction Progress Monitoring		(17%)	**(83%)**			(34%)	**(66%)**
	Space-Time Logistics Management	(17%)	**(66%)**	(17%)			**(66%)**	(34%)
	Change Management		(34%)	**(66%)**			**(50%)**	**(50%)**
Post-Construction	Commissioning & Handover		(33%)	**(67%)**		(33%)		**(67%)**

Table S9. Perceived Process Bottlenecks Based on Industry Expert Input

Phase	Process	Synonyms
Preconstruction	Contract management	Contracting, contract values
	Procurement management / Procure-to-pay	Materials procurement, equipment procurement
	Order-to-cash	Subcontract, AR (Accounts receivable)
	Design approval/Authoring	CAD (Computer Aided Design) or VDC (Virtual Design & Construction)
	Document control management	SharePoint, Procore (technology to support process)
	Submittal & Transmittal approval	Technical submissions
	Progress payments for design	Invoicing, payment terms
Construction	Resource utilization management	Resource management, Labour management, (Labor, material, equipment and schedule)
	Material requisition/inventory/tracking management	Material purchase orders
	Change Management	It can involve change events, change orders, prime contract change orders, Site Instructions (SI), contemplated change notices (CCN), change requests (CR)
	Construction progress monitoring	Work in Progress (WIP); Progress Tracking
	Site logistic and space management	Staging area, laydown areas
	Quality assurance & Quality control	Deficiencies Assessment & Compliance Assurance

Table S10. Analysis of previous studies on RFI Management

Authors	Year	Location	Construction Project Type	Method(s) / Techniques	Main Metrics or KPI(s)	Main Focus/Goal
Afzal et al.	2023	Australia	Civil & Residential	NLP LDA	Silency & Relevance (Based on frequencies & probabilities)	RFI Classification
Yilmaz and Ergen	2024	Greece	Infrastructure (Airport project)	NLP & LDA K-means clustering Matrix Factorization	Silency , Relevance & Euclidean distance Matrix coefficients	RFI Classification
Hanna et al.	2012	USA	Infrastructure (Highways)	Bootstraping statistical analysis	No.of RFIs: % of On-time RFIs	Define general performance benchmarks
Sheela	2015	Canada	Institutional & Commercial	Inductive and deductive content analysis (manual) & EDA	Various Visual metrics No.of RFIs: RFI latency	Structuring and Visualization of RFI Data
Ozogul and Ergen	2024	Greece	Not Specified	NLP Naïve Bayes (NB) & K-nearest neighbor (KNN)	Precision Recall F1 Score	Automated extraction of RFI Metadata
Panahi et al.	2023	USA	Healthcare Building Projects	NLP, YOLO 5, OCR, GPT Keyword and semantic-search similartiy	Cosine Similarity No. of RFIs.	RFI recommender system to streamline pre-construction design reviews
Mao et al.	2006	USA	Building Projects	Content analysis (manual) IFC-based metadata model	N/A	RFI interoperability through automated extraction of RFI metadata
Kim et al.	2022	USA	Civil Projects	Statitical Analysis (t-test and one-way Anova)	t-value p-value	Comparative cost growth due to RFIs on DBB and DB projects
Adamtey	2020	USA	Various (Industrial, commercial & civil)	Statitical Analysis (one-way Anova)	p-value Time overrun (TOR) Cost Overrun (COR)	Time and Cost overruns on PDB and DB projects
Aibinu et al.	2020	Autralia	Mixed-used residential and commercial	Statitical Analysis (one-way Anova)	p-value RFI frequency RFI Turnaround time	Impact of project characteristics on RFI frequency and RFI turnaround times

Table S11. Analysis of previous studies on RFI Management (Continuation)

Authors	Year	Location	Construction Project Type	Method(s) / Techniques	Main Metrics or KPI(s)	Main Focus/Goal
Ibrahim et al.	2020	USA	Industrial, instrastructure, buildings	Statitical Analysis (Anova)	p-value Kendall's tau-b test	Quantitative impact of Out-Of-Sequence (OOS) work on Project Performance
Das et al.	2022	China	Not Specified	Blockchain-based smart contract logic	latency and scalability	Security enhancement of EDMS (decentralized) and traceability of approval workflows
Erri Pradeep et al.	2021	New Zeland	Not Specified	Blockchain-based smart contract logic	Not Specified	Security and traceability of information exchanges
Golzarpoor et al.	2016 2018	Canada	Oil & Gas projects	Worflow inheritance-rules conformance checking (IFP ontology-based)	Rule-based checking	Automated Process Conformance & Interoperability
Papajohn and El Asmar	2020	USA	Highway projects	Descrete Event Simulation	RFI Response time	Assessement of RFI Response Time under different Project Delivery Systems
Mohamed et al.	1999	England	Residential Building projects	System Dynamics Simulation	RFI Cycle time RFI Cost	Quantify Time & Cost Impacts of the RFI Management Process
Kouhestani and Nik-Bakht	2020	Canada	Not Specified	Inductive Visual Miner Fuzzy Miner SNA	Processing time	Design authoring process (process discovery, conformance checking, and bottleneck analysis)
Pan and Zhang	2021	Singapore	Residential	Inductive Visual Miner Fuzzy Miner SNA	SNA Various Metrics Fitness; Precision;	Organizational Performance
Wang et al.	2024	China	Civil Highway Projects	Inductive Visual Miner Fuzzy Miner Process Variant Analysis	Fitness; Precision; time bottlenecks	Highway construction process (process discovery, conformance checking, and bottleneck analysis)
Martinez Lagunas and Nik-Bakht (This Study)	2025	Canada	Various Building Projects (Industrial, Residential, Healthcare, Institutional, Commercial)	LPMM Inductive Visual Miner Fuzzy Miner SNA NLP & Process Variant Analysis	WACE (Cost Efficiency) WAPE (Time Efficiency) SNA Various Metrics Fitness; Precision; F-score	Automated Process Performance Monitoring (Process Efficiency)

Appendix S10. Procore-based Glossary of Financial Terms [199]

A. Contractual and Financial Provisions

Allowance: A predefined monetary amount included in the original contract value to cover specified work items whose detailed scope or specifications were not fully defined at the time of contract execution.

Contingency: A reserved portion of the contract sum intended to address unforeseen costs arising from uncertain or unknown conditions not explicitly defined in the contract documents.

Guaranteed Maximum Price (GMP): A contractual arrangement in which the contractor is reimbursed for actual costs incurred plus a fixed fee, subject to an agreed upper cost limit beyond which the contractor assumes financial risk.

Lump Sum Contract: A contracting method in which the contractor agrees to deliver the defined scope of work for a fixed total price, inclusive of overhead, profit, and all direct and indirect costs.

Cost Plus Contract: A contract type where the owner compensates the contractor for actual project costs in addition to an agreed markup or fee.

Liquidated Damages: Pre-agreed financial penalties, typically assessed on a daily basis, payable by the contractor to the owner for failure to complete the work within the contractually specified time.

Retainage: A contractual practice whereby a percentage of each payment is withheld until specified milestones or substantial completion are achieved, serving as a financial assurance of performance.

B. Change Management and Cost Control

Request for Information (RFI): The term Request for Information (RFI) refers to a formal business process initiated by a project participant, such as a general contractor, subcontractor, or supplier, to seek clarification or formally communicate issues related to project documentation, design, or execution. RFIs are reviewed and responded to by the project owner's representative or the architect, and the resulting responses may influence the project's scope of work and, when cost or contractual implications arise, require formal approval.

Change Event: A documented occurrence that alters the project's scope, schedule, cost, or execution conditions and may lead to a formal contract modification.

Potential Change Order (PCO): A preliminary cost and scope assessment prepared by the general contractor to evaluate the potential financial impact of a change before formal approval.

Change Order (CO): A formally approved modification to the original contract that revises the agreed scope, schedule, and/or cost.

Prime Contract Change Order (PCCO): A change order that directly modifies the prime contract between the owner and the general contractor, often aggregating multiple PCOs or Change Order Requests.

Commitment Change Order (CCO): A contract amendment that adjusts the financial commitments between the general contractor and subcontractors or vendors.

Change Order Request (COR): A structured submission consolidating one or more potential changes into a formal request for owner review and approval.

Change Directive: An owner-issued instruction requiring immediate execution of changed work prior to formal agreement on cost or schedule adjustments.

Change Notice: A formal communication informing stakeholders of a proposed or pending change that may affect contractual obligations.

Cost Code: A standardized alphanumeric identifier used to categorize and track specific work activities and associated costs within a project.

Cost Type: A classification scheme used to distinguish categories of expenditures such as labor, materials, equipment, subcontracted work, or professional services.

Enterprise Resource Planning (ERP): Integrated business software systems used to manage organizational data across accounting, procurement, construction operations, and financial reporting.

D. Contracts, Procurement, and Payments

Prime Contract: The primary legal agreement between the project owner and the general contractor governing project delivery responsibilities and obligations.

Request for Quote (RFQ): A formal solicitation inviting vendors or subcontractors to submit pricing proposals for specific goods or services, often related to change events.

Invoice: An itemized financial document requesting payment for completed work or delivered services.

Owner Invoice: A billing statement issued to the project owner detailing approved costs incurred during a billing period.

Progress Payment: A partial payment issued periodically based on verified completion of portions of the contracted work.

E. Estimation, Scheduling, and Value Allocation

Rough Order of Magnitude (ROM): A high-level cost estimate used during early project phases to provide an approximate indication of potential project costs under high uncertainty.

Schedule of Values (SOV): A detailed breakdown of the total contract amount into individual cost items used to support progress measurement and payment certification.

Scope of Work (SOW): A detailed description of the tasks, deliverables, and responsibilities required under a contract or subcontract.