Semi-Automated Methods for Measuring Practice Conformance for Capital Projects

by

Seokyoung Kang

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Civil Engineering

Waterloo, Ontario, Canada, 2020

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	**Brenda McCabe** Professor, Department of Civil & Mineral Engineering, University of Toronto
Supervisor	**Carl T. Haas** Professor, Department of Civil & Environmental Engineering, University of Waterloo
Internal Members	**Scott Walbridge** Associate Professor, Department of Civil & Environmental Engineering, University of Waterloo
	Chris Bachmann Assistant Professor, Department of Civil & Environmental Engineering, University of Waterloo
Internal-external Member	**Derek Rayside** Associate Professor, Department of Electrical and Computer Engineering, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contribution

Seokyoung Kang was the sole author of Chapter 1, 2, 5, 6, and 7 which were written under the supervision of Dr. Carl T. Haas and were not written for publication. This thesis consists in part of two manuscripts written for publication. Exceptions to sole authorship of material are as follows:

Research presented in Chapter 3:

This research was conducted at the University of Waterloo by Seokyoung Kang under the supervision of Dr. Carl T. Haas. Dr. Carl T. Haas, Dr. Behrooz Golzarpoor, Dr. Ekin Eray, Dr. Mahdi Safa, Dr. Derek Rayside, Matthew Weston, Joel Gray, members of Construction Owners Association of Alberta (COAA), and Seokyoung Kang contributed to model design, survey participant recruitment, and statistical analysis presented in Chapter 3. Dr. Behrooz Golzarpoor, Dr. Carl T. Haas, Dr. Derek Rayside, Seokyoung Kang, and Matthew Weston co-authored a journal paper. Seokyoung Kang contributed by reviewing past studies and suggesting potential solutions. Seokyoung Kang and Dr. Carl T. Haas co-authored a conference paper. Seokyoung Kang contributed by searching for existing commercially available tools, and by testing and comparing their functionalities. Seokyoung Kang, Dr. Carl T. Haas, and Dr. Behrooz Golzarpoor co-authored a conference paper. Seokyoung Kang contributed by developing a framework and providing examples. Dr. Behrooz Golzarpoor, Dr. Carl T. Haas, Joel Gray, Dr. Derek Rayside, and Seokyoung Kang co-authored a conference paper. Seokyoung Kang contributed by outlining the conformance model and creating connections between two concepts-interoperability and conformance.

Citation:

Golzarpoor, B., Haas, C. T., Rayside, D., Kang, S., & Weston, M. (2018). Improving construction industry process interoperability with Industry Foundation Processes (IFP). Advanced Engineering Informatics, 38, 555-568.

Kang, S., & Haas, C. T. (2018). Evaluating Artificial Intelligence Tools for Automated Practice Conformance Checking. *2018 Proceedings of the 35th ISARC*, *35*, 110–117. https://doi.org/10.22260/ISARC2018/0015

Kang, S., Haas, C. T., & Golzarpoor, B. (2017). Framework for Measuring Process Interoperability in Construction Management. *Annual Conference of the Canadian Society for Civil engineering 2017*, CON070-1-10

Golzarpoor, B., Haas, C. T., Gray, J., Rayside, D., & Kang, S. Industry Foundation Processes (IFP) – A Unique Approach to Improve Process Conformance and Interoperability. *Annual Conference of the Canadian Society for Civil engineering 2017*, CON119-1-10

Research presented in Chapter 4:

This research was conducted at the University of Waterloo by Seokyoung Kang under the supervision of Dr. Carl T. Haas. Liyuan Chen, Dr. Karimidorabati, and Seokyoung Kang contributed to create the framework and design methodology. Seokyoung Kang contributed by finding patterns in the event log and devising the solution for each case.

Citation:

Chen, L., Kang, S., Karimidorabati, S., & Haas, C. T. (2019). Improving the Quality of Event Logs in the Construction Industry for Process Mining. *2019 Proceedings of the 36th ISARC*, *36*, 804–811. https://doi.org/10.22260/ISARC2019/0108

Abstract

The goal of this thesis is to explore semi-automated methods for measuring practice conformance for capital projects. Thorough measurement of practice conformance for capital projects typically requires manual audits. Surveys that may assist can often be subjective, non-repeatable and unverifiable, since they are self-reported. However, some of the tasks assigned to auditors are also non-repeatable, and they may be costly, time-consuming, tedious, and error-prone. Tools for assisting practice conformance measurements are in high demand in the construction domain. In response, various information technology-based and web deployed Benchmarking and Metrics (BM&M) programs have been introduced to reduce time and costs, to assist in providing repeatable and accurate results, and to increase efficiency and productivity of reporters and auditors. Moreover, moves toward automated practice conformance measurement are expected to reduce time and cost. Past studies have also resulted in significant advances in data mining, natural language processing, machine learning, computer vision and other artificial intelligence-based approaches toward complete automation, but technical limitations exist that constrain complete automation or make it impractical. An approach is needed to support practical, net beneficial, incremental steps toward automation of practice conformance measurement for capital projects that would assist capital project participants to improve project performance over time. To address this need, a new approach is proposed in this thesis. Additionally, a framework to beneficially increase automation is presented. Toolsets are explored that may make practice conformance measurement cheaper, faster, easier, repeatable, and more accurate for capital project participants. This framework and the toolsets are validated through the development of a practice conformance model, case studies on real project data, and application experiments. It is concluded that the proposed semi-automated framework for measuring practice conformance for capital projects is practical to implement in the near term. These results provide a basis on which capital project participants can implement efficacious practice conformance measurement to support capital project performance improvement programs.

Acknowledgements

I am deeply thankful for my supervisor, Carl T. Haas. Not only because he taught me how to write papers and be a scholar, but he also taught me how to voice my opinion, make logical statements, and communicate well with others. Despite all my weaknesses, he always encouraged me with compliments and guided me to the best way. I am also thankful for my committee members for their advice. I have a deeper understanding of this field because of all the questions and comments provided from them.

This research was not possible without the help of industrial partners who provided data and willingly participated in a survey. I was privileged to have support from COAA, CII, and Coreworx. I also would like to express my gratitude to my colleagues who shared ideas and opinions. I was able to grow and mature intellectually by walking along the journey with them. I cannot name every colleague and advisor, but, they will be remembered.

I would like to thank my friends who were there for me when I needed it. Not only emotional support but practical advice from them helped me achieve my goal. I am grateful to Charmian Wenger, Florence Tsang, Olivia Merritt, and Janelle Resch for editing the grammar of my thesis.

The supports from my family are truly a blessing. I am thankful to mom (Jodie Cho) and dad (Jeff Kang) in Korea for giving me the best opportunities and supporting all my decisions. I am also very grateful to my mother and father in law (Dorman & Sheryl Martin) for their genuine love and care for me. So thankful to my brothers (Jason Kang & William Martin) for their encouragement.

I want to thank my husband, James Martin, for being the one who always reminds me of who I am and what I can do. I can be myself because he is by my side, because of his belief in me, and because of his everlasting love towards me. Lastly, I thank God for all He has done in and through me.

Dedication

To my husband,

James Martin

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

As a part of an operational audit, practice conformance needs to be measured in the construction industry. Third-party audits have advantages in that they are more objective and unbiased; however, the audit fee is expensive due to numerous documents, complex workflows, and various interviews that an auditor has to go through. One way to circumvent the high audit fee is to automate some parts of the practice conformance measurement. This thesis proposes that practice conformance measurements can be semi-automated with the appropriate model, framework, and toolsets. This thesis proposes a simple practice conformance model, a consistent and repeatable framework, and quick, inexpensive, and accurate toolsets. Companies can select parts of the proposed framework and suggested toolsets to apply to their particular project to increase audit productivity. A model is established based on the literature review and current practices, a framework is tested, and toolsets are evaluated through functional demonstration, case studies and application experiments.

1.1 Research Background and Needs

Most construction companies are familiar with construction audits and financial audits. A construction audit is an analysis of the costs incurred and actions taken for a specific construction project. Contracts granted to contractors, the price paid, overhead costs allowed for reimbursement, change orders, and the timeliness of completion are some of the activities that auditors inspect. A financial audit is an analysis conducted by a third party that increases the credibility of financial statements and reduces risks for stakeholders. However, there are more than construction and financial audits in the audit category. A compliance audit is an examination of the policies and procedures of a company to see if it follows internal or regulatory guidelines. An operational audit requires a detailed analysis of the goals, planning processes, procedures, and results of the operations of a business. There are also information systems audits, investigative audits, and tax audits. Among these various types of audits, practice conformance measurements, as a part of operational audits, are the focus of this thesis. The scope is further narrowed to those activities, which can possibly be semi-automated to meet the needs of construction companies and auditors.

Such audits can be executed by first, second or third parties. A first-party audit is an internal audit conducted by auditors who are employed by the company being audited but who have no interest in the audit results of the area being audited. A second-party audit is completed by an external party or a contracted company on behalf of a customer on that customer's operations. A third-party audit is completed by an audit organization independent of the customer-company relationship and is free of any conflict of interest. Independence of the audit organization is a key component of a third-party audit.

Operational audits are typically internal audits executed by internal or outsourced staff. Internal audit should not have any operational accountability or perform functions that would be subject to subsequent internal audit review as opposed to financial audits which have stricter regulations. Therefore, operational audits generally do not face regulatory hurdles or have limitations in using technologies. Audit process automation is encouraged (Pelland, 2017) and operational audits are recommended to adapt its methodologies to increasingly utilize technology in the execution of audits (KPMG, 2020).

According to a U.S. Securities and Exchange Commission (SEC) filing, the median annual audit fee for over 6,000 public companies is more than $520K in 2016 at an hourly rate of $216 (Pelland, 2017). The reason for the high hourly rate is that experienced professionals are required to go through thousands of pages of documents, complex workflows, and personal interviews.

Companies are suffering from the high cost and desire to optimize the audit process. As well, auditors do not enjoy wasting time on tedious work. Auditors typically desire to contribute to solving the main issue, such as improving processes or implementations. However, currently, it takes time just to decide where and what to focus on for auditors. Moreover, tedious tasks take many of the hours needed to complete the audit. Auditors who need to do critical tasks for companies are often wasting their time on tasks with little purpose. Companies report pressures on their audit fees. Yet the same companies to some extent still partially rely on self-reported surveys. To mitigate audit fees and avoid self-reported surveys, companies strive to increase collaboration with auditors, to focus on key audit areas, and to automate internal controls processes (Pelland, 2017).

Another way to reduce the audit fee is to automate some parts of the audit process. Automation enables companies to negotiate rates that are fair for both auditors and the audited companies. If tedious tasks previously completed by auditors are accomplished in automated ways, auditors can be more efficient. With semi-automation, companies may pay less in total for an audit that still includes the more critical tasks that are provided by auditors, and auditors may still be satisfied by their contributions to the company.

1.2 Research Posits

The following posits, which will help establish the research objectives, are proposed:

- Practice conformance measurement, as part of an audit, may be semi-automated.
- The model proposed by this research is simple; the framework is consistent and repeatable and; the toolsets are inexpensive, quick, and accurate enough to assist auditors.
- The model, framework, and toolsets proposed by this study can be selectively applied to increase audit productivity and accuracy.

1.3 Research Objectives

The primary goal of this thesis is to explore semi-automated methods for measuring practice conformance for capital projects. The research objectives of this thesis follow from that goal and the posits proposed in the previous section. The objectives, therefore, are to:

1. Establish a capital project practice conformance framework and model based on a critical review of the literature and current practice.

2. Discover and examine methods for the automation of capital project practice conformance measurement, and evaluate their potential in terms of adequate accuracy, ease of use, cost, and acceptability by project participants.

3. Test two or more high potential methods experimentally and demonstrate their potential to assist auditors and audited companies to increase audit productivity and accuracy.

4. Critically evaluate if the framework and toolsets proposed in this research are repeatable, consistent, inexpensive, quick, and accurate enough to assist auditors.

The scope of these objectives is constrained, as explained in the following section, and then the methods for achieving these objectives are described in Section 1.5.

1.4 Research Scope

To accomplish the objectives, scope is defined as highlighted in the Venn-diagram (Figure 1-1). With current technology, limitations exist to achieve full automation; therefore, semi-automation is the scope. Only written documents and clearly defined processes are addressed. Practices completed by experience without any documents or clear processes are, therefore, not the scope. Capital projects are typically large, complex and capital intensive. Owners of capital projects can both invest in this research and make the best out of its findings.

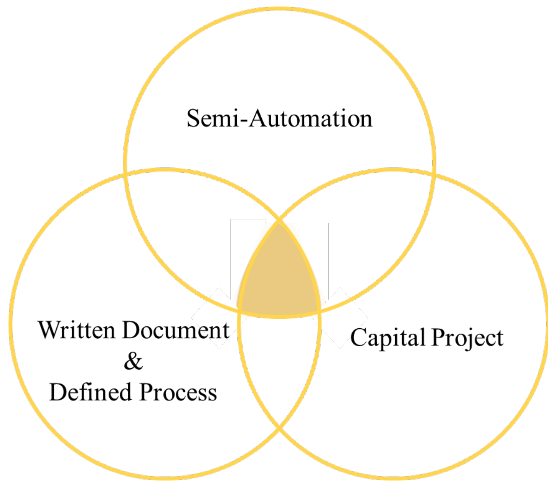


Figure 1-1: Research Scope

This thesis provides an explanation of a practice conformance model and a conformance measurement framework and suggests toolsets to improve the productivity and accuracy of an audit. Among various ways to make audits cheaper, quicker, and easier, this study focuses on the automation of the practice conformance measurement. Full automation is not realistic given technological limitations and insufficient data. However, semi-automation potentials for practice conformance measurement of documents and workflows are identified in this study. Commercialized tools are examined, and detailed illustrations of the ways to achieve semi-automation are provided throughout the thesis.

Measuring practice conformance to a guideline document and a benchmark workflow is especially beneficial to acknowledge the influence of the practice guideline and the benchmark workflow. Nevertheless, some companies do not have documented practice guidelines or structured workflow in house but recall the processes from memory and perform them without a problem. This is typically the case for small simple projects. These cases are not within the scope of this study, as large capital projects involve more complex structures and complicated processes. The capital projects that require documented practice guidelines and structured benchmark workflows are the focus of this study as this research will have a greater impact on them.

The functional demonstrations and case studies in this thesis are based on current practices in capital projects; specifically, the data collected to establish the conformance model are based on oil and gas industry. The need for this study and the capacity to take advantage of it are the main criteria for selecting the scope. However, the model, framework, and toolsets for the semi-automated practice conformance measurement can be used in various scales of construction projects and possibly in other industry domains, such as IT, manufacturing, health, and training that utilize practice guidelines and benchmark workflows to improve their current practice.

1.5 Research Methodology

In the preliminary stage, problems are investigated, and relevant literature is reviewed. Then, based on the literature review, an initial practice conformance model is synthesized. The practice conformance components which constitute a conformance model are derived inductively, supported by expert consultations and meetings, while practice conformance elements are validated through an in-depth survey based upon grounded theory research methodology. To distinguish between practice conformance elements and practice conformance components, a component is a broader scope that is required to achieve conformance. An element represents a tangible subset of data which provides a measurable form of evidence to conformance. An expert refers to a professional with more than 30-year experience in one industry sector whom can represent the industry perspective. These professionals participated in the validation survey which aided in developing the conformance elements. Practice conformance components and practice conformance elements are further explained in later chapters. The final practice conformance model is then established. The potential of the semi-automated practice conformance measurement methods is validated through functional demonstrations of the framework and toolsets. Case studies are conducted with real-data collected from the construction industry. Performance experiments are also conducted. The results validate that the practice conformance framework is useful, and they are documented in the thesis. A summary of the research methodology is illustrated in Figure 1-2.

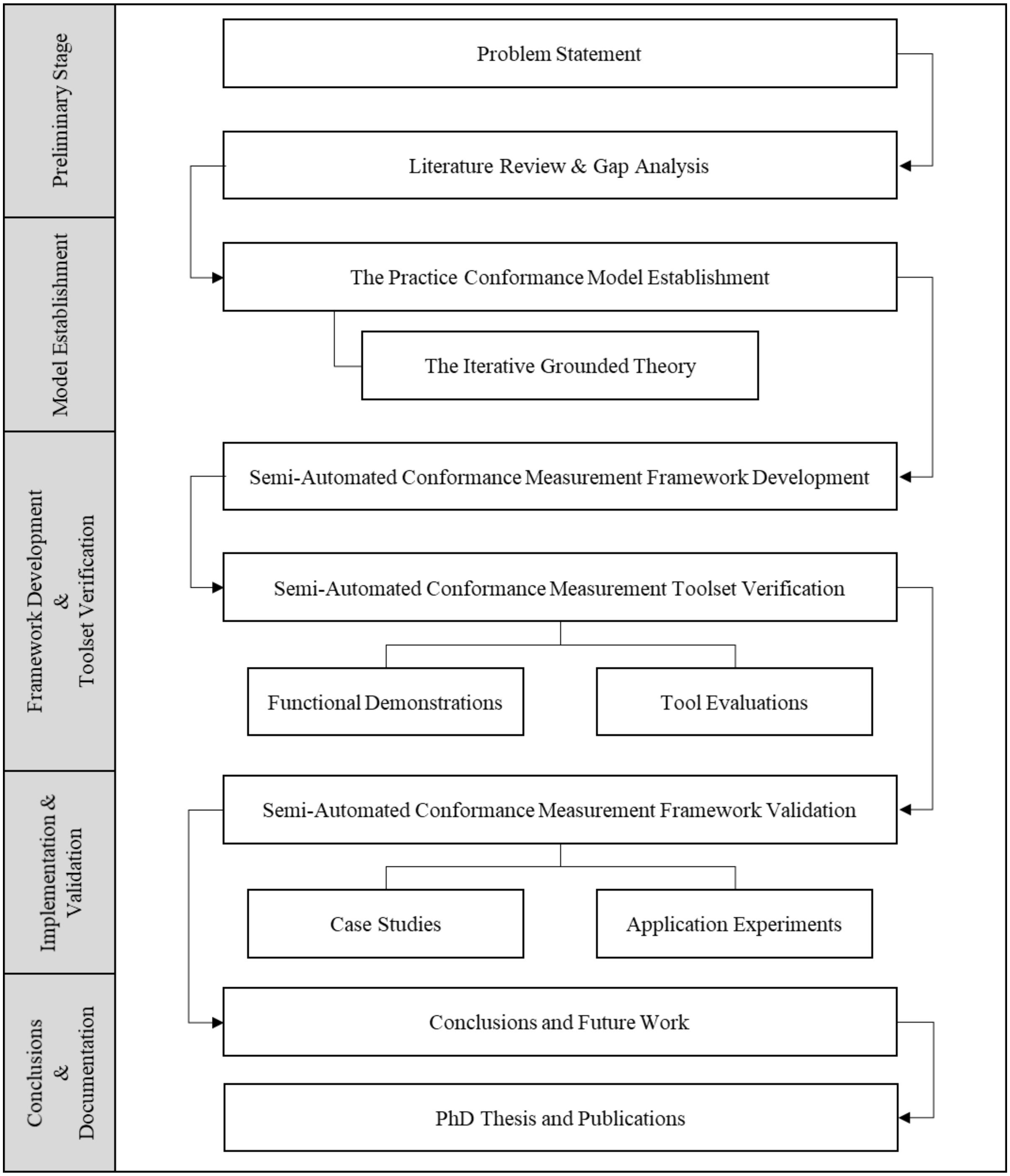


Figure 1-2: Research Methodology

1.6 Structure of the Thesis

This thesis consists of seven chapters. In Chapter 1, background and needs for the research are described. Research posits, objectives, and scope are defined along with a summary of the research methodology. In Chapter 2, literature regarding conformance and other relevant topics to the thesis is reviewed, and a knowledge gap is identified. In Chapter 3, the practice conformance model is established which includes practice components, measurement methods, and practice conformance elements. In Chapter 4, the practice conformance measurement framework is developed for document and workflow conformance. In Chapter 5, commercialized toolsets for assisting the semi-automated framework are explored and verified. Functional demonstrations are presented. In Chapter 6, case studies are conducted with the selected methods to validate the research posits and to complete the research objectives. Lastly, in Chapter 7, the thesis concludes with a summary, contributions, limitations, and suggested future work.

Chapter 2 Literature Review

Practice conformance in various industry sectors such as healthcare, manufacturing, and banking has been significantly improved with automation in analyzing processes and in evaluating the implementation of such processes. In the construction industry, however, practice conformance is measured manually by self-reported surveys or audits, which require the use of experts. Considering that most capital projects are managed by hundreds and thousands of pages of documents and sets of complex workflows, automation in the measurement of practice conformance is perceived to be vital.

Past studies have proved the importance of automating the measurement of practice conformance, but, due to the complexity or pioneer stage of their approaches, it has not been practical for the companies to implement theoretical research, such as Automated Compliance Checking (ACC) or process mining. Moreover, the implementation entails lower accuracy because of technical limitations and data collection issues. While Benchmarking and Metrics (BM&M) program-related literature is more practical, BM&M studies also have limitations because they are mostly based on self-reported surveys. This thesis improves data collection methods by utilizing machine recorded data rather than human recorded data. Though past studies argue that full automation has been achieved, due to several exceptions, their approaches are hard to apply to the real-world. Thus, to increase accuracy and to be more industry-applicable, this study supports the semi-automation framework rather than full automation.

2.1 Compulsory vs. Non-Compulsory Practice Guidelines

Compulsory and non-compulsory practice guidelines must first be defined because they are used to develop a framework in the later chapters. A successful capital project requires both practice guidelines and the judgment of professionals. A practice guideline is an operational manual that is used to enhance the quality of a product, process, or service. Practice guidelines might exist in various formats, including an electronic format. Depending on the force of law, practice guidelines are divided into compulsory guidelines and non-compulsory guidelines. Compulsory guidelines involve laws, policies, and codes that have the force of law, while non-compulsory guidelines involve best practices, benchmark workflows, and standard procedures that do not have the force of law. With these guidelines, professionals extract information, absorb it into their knowledge, and apply it in order to make wise decisions based on their experiences (Ackoff, 1989). Both compulsory and non-compulsory practice guidelines can be used to measure conformance. They are intended for companies to conform (or adhere) to so that each company can achieve their ultimate goals. Ultimate goals not only include company objectives, but also the visions and values of the community or government, such as the safety of workers and environmental issues.

Legislation refers to written laws, often referred to as acts or statutes, that are enacted by parliament, the legislative arm of government. Regulations are a form of law, sometimes referred to as subordinate legislation, that define the application and enforcement of legislation. The term policy outlines the goals that a government ministry desires to achieve and the methods and principles it will use to achieve them. The policy document is not a law but often identifies new laws needed to achieve its goals. A code is a set of rules that experts in the field recommend people to follow. Although it is not a law, it can be adopted into law. Code refers to what needs to be done but does not explain how it should be done.

When it comes to non-compulsory practice guidelines, a standard is a document approved through consensus by a recognized body with which conformance is not compulsory. The recognized body (or an institution) typically provides rules, related processes, and production methods for common use. Standards are embodied in several forms, such as the definition of terms, specification of design and construction, detailing of procedures, or performance criteria against which a product or process can be measured. When a standard has been adopted by governmental bodies and has the force of law, it becomes a code. A standard also becomes a code when it has been incorporated into a business contract. Both standards and codes establish technical or engineering requirements for products, methods, practices, or operations. Another non-compulsory practice guideline is a best practice. A best practice is defined as a process or method that, when executed effectively, leads to enhanced project performance. The construction industry is relatively reluctant to adopt new practices because of their scales and risks (Rahman, 2014). While some companies are hesitant when it comes to conforming to non-compulsory practice guidelines like best practices, best practices can diminish potential risks because they are established from numerous past cases.

When compared to laws or regulatory statements, best practice conformance means doing beyond what one must do. The intent is to perform better by following the processes described in the best practices. Best practices might take longer or incur higher costs in the beginning due to the learning curve (Randolph et al., 1986). However, they have positive influences on the overall project. Although compulsory practice guidelines might be enough for some projects, it is advised to consider the best practices for large and complex capital projects (Cha & O'Connor, 2005). Best practices are not intended to increase work or restrict progress, but are intended to avoid risk and rework, driving better project outcomes by applying best practices selectively and effectively.

2.2 Capital Project Benchmarking and Metrics (BM&M) Programs and Data Collection Practices

A Benchmarking and Metrics (BM&M) program and operational audit program share some common goals and methods, one being measuring conformance. Thus, past studies in relation to capital project benchmarking and metrics programs are reviewed. The benefit of the data collection approach discussed in this section is that significantly more data is available than what can be afforded in a site-visit approach. However, responses are self-reported; therefore, are subject to some level of interpretation and might be less credible than the results of site visits to each project and company. Additionally, the data is manually recorded which increases the risk of human errors.

Benchmarking and metrics programs aim to measure and assess capital project performance and to find best practices among similar projects (Shan et al., 2011; Zhai et al., 2009). Once best practices are defined, they are implemented on future projects, and the productivities of these projects are measured. To assess the conformance and performance of the best practices, measurement criteria must be defined. Then, suitable metrics must be derived or selected, and data must be collected and analyzed.

The use of effective conformance and performance metrics has been key to the successful benchmarking of capital projects. Several benchmarking studies have developed measurements to evaluate various performance outcomes (Rankin et al., 2008; Yun et al., 2016). Once a critical mass of data is acquired, statistically significant results can be reported with BM&M programs, and a cycle of positive reinforcement can begin.

Designing BM&M programs is not simple. Rankin et al. (2008) designed a timeline for performance metrics. The construction timeline is divided into six phases with seven points. To develop metrics, the points are used in the definitions, and formulas are created accordingly. Cost, time, and scope metrics are quantified using estimated and actual cost and time. Safety metrics are quantified using working hours and number of incidents. Quality satisfaction is measured using ratings.

Although the metrics by Rankin et al. (2008) were understood, they were complex and time-consuming to record (Nasir et al., 2012). Nasir et al. (2012) improved a BM&M program for construction performance and productivity improvement with the input-output diagram. Input refers to materials, personnel, equipment, management, and money, while output refers to the number of housing units, kilometers of highway, and cubic meters of concrete. Conditions vary according to the complexity of the design, type of construction, and environmental factors. Objectives are in relation to cost, time, quality, and/or safety. With a scoring system, three classes of metrics were developed: project performance, labor productivity, and management practices. Details to develop and implement BM&M programs are further explained.

Nasir et al. (2012) described how data collection programs are typically structured. Phase One establishes the BM&M program by setting preliminary productivity metrics, setting international comparison methods, developing a framework of practices, setting data collection protocols and tools, developing an initial data repository, and evaluating program components with experts. In Phase Two, pilot data is collected by prototyping the data collection tools and processes, developing communication~~s~~ materials, conducting stakeholder regional workshops, and revising the data collection tools and communication~~s~~ materials. Phase Three extends the pilot phase and collects full data by collecting from multiple projects, analyzes that data, and establishes sustainable benchmarking programs. Nasir et al. (2012) simplified the original approach and focus on seven specific topics. Yes-or-no questions were perceived to make data collection easier to understand and quicker to respond. To collect the data for these BM&M programs, literature review, face-to-face meetings, conference calls, workshops with industry stakeholders, industry advisory group reviews, and kick-off took place. Though there were various approaches, self-reported responses were unavoidable.

Zhai et al. (2009) used self-reported surveys by project participants for project process automation practice use-level assessment. To begin the study, literature reviews, acquisition of documentation from owners and contractors, and a series of workshops were conducted. Automation and integration use-levels for thirteen work functions (i.e., practices) were defined. The Construction Industry Institute (CII) five-point scale was used for the survey. Metrics were used to derive project-level automation and integration indices and to create a hypothesis between low and high automation/integration. To facilitate statistical analysis, the scores were grouped into high or low project use-level of automation and integration. The T-test was conducted to find the significance of differences.

Shan et al. (2011) defined productivity, standard labor productivity metrics, and use-index (i.e., practice conformance rate) to measure how well the best practices are implemented. The work of the sample projects was relatively similar, yet samples were not under controlled environments. The self-reported survey methodologies adopted by the BM&M program database of the Construction Industry Institute (CII) were used for the analysis.

To measure conformance to best practices, a level of implementation of productivity practices (PIL) index was developed (Caldas et al. 2015). The best productivity practice implementation index (BPPII) is defined in terms of categories, sections, and elements. Although the weights were determined by the survey and each level was defined, data collection depended on self-reported survey forms.

Studies regarding the definitions and frameworks that are used to develop BM&M programs, are typically completed by self-reported survey questionnaires and manual recordings. While the past studies used self-reported surveys, it is possible for auditors (or a third party) to complete the questionnaires with a site visit; however, this is perceived as costly and time-consuming. Moreover, manual recording still remains an issue. Thus, there is a demand to balance self-reported surveys and expensive audits. Table 2-1 summarizes the strengths and weaknesses of past important foundational studies regarding BM&M programs.

Table 2-1: Summary of the Literature Review on BM&M Activities Related to Conformance Measurement

References	Keywords	Strength				Weakness			Notes
		D/M	F/P	DA	LDC	S	NQ	MRD	
Rankin et al. (2008)	Phase points	X	X			X	X	X	Subjective results based on estimates
Zhai et al. (2009)	Integration; Automation	X		X	X	X		X	Definitions are vague (e.g., little utilization)
Shan et al. (2011)	Mechanical trades			X	X	X		X	
Nasir et al. (2012)	Labor productivity		X			X		X	Takes a long time to build
Caldas et al. (2015)	Best Practice Implementation Index (BPPII)	X	X			X		X	
Yun et al. (2016)	Key Performance Indicators (KPIs)	X		X		X		X	Depends on the forecast

D/M: Definition/Metrics; F/P: Functional demonstration/Pilot study; DA: Detailed Analysis; LDC: Large Data Collection; S: Subjective; NQ: Non-Quantifiable; MR: Manually Recorded Data

Many of the past studies regarding BM&M created new definitions and/or metrics. They were mostly derived inductively. Some of them were applied in functional demonstrations or pilot studies, while others were used for detailed analysis with a large data collection. However, the results were relatively subjective because the data was collected through self-reported surveys. The results were also prone to error because the surveys were recorded manually. Each study mentioned in Table 2-1 has other limitations, including metrics that were not quantifiable, vague definitions, long durations to build the data collection, or a dependency on a forecast that can vary in time.

2.3 Automated Compliance Checking (ACC)

Automated compliance checking (ACC) and document conformance measurement proposed in this study focuses on both texts and semantics. Because document conformance embraces ACC (even though the approach might differ), it is worth exploring backgrounds and the latest findings.

Yurchyshyna and Zarli (2009) conceptually demonstrated how to capitalize and organize knowledge for conformance checking in construction using automated ontology-based reasoning. This preprocessing stage checks the construction process against technical norms. Yurchyshyna et al. (2010) further developed the previous studies by adding technical details and presenting the conceptual conformance checking model.

Eastman et al. (2009) reviewed the rule checking process, mainly focusing on geometric rules. Some software packages for rule-based building model checkers were investigated. According to Eastman et al. (2009), rule checking has limited reporting capabilities, error reporting, and quick corrections. Since then, Salama and El-Gohary (2011) emphasized the need for automated compliance checking (ACC) and present new frameworks and methodologies using deontic logic. Salama and El-Gohary (2013) further developed a detailed framework and establish hierarchies and relationships among documents. Salama and El-Gohary (2016) used machine learning to classify texts and extract information.

Zhang and El-Gohary (2012) simplified the human language by dissecting it and applied existing grammatical rules. Zhang and El-Gohary (2013) complemented previous studies with information transformation based on semantic mapping rules, conflict resolution rules, and consume and generation mechanisms. Zhang and El-Gohary (2015) then specified information extraction by applying tokenization, sentence splitting, morphological analysis, de-hyphenation, and Part-of-Speech (POS) tagging. Zhang and El-Gohary (2016) emphasized information extraction from regulatory documents to match patterns. Zhang and El-Gohary (2017a) applied logic reasoning to analyze the texts and to create relationships among words. Zhou and El-Gohary (2017) built upon and summarize previous studies by elaborating an automatic information extraction procedure in six steps. However, moving beyond conceptual quantitative comparison is yet to be achieved. Additionally, for these approaches, ACC is not possible without significant amounts of data.

Building information modeling (BIM) is currently receiving more attention in relation to compliance checking. In contrast to extracting texts for cross-checking, Zhang and El-Gohary (2017b) attempted to automatically extract design information from building information models and proceeded to check compliance with regulatory information using semantic-based logic reasoning algorithms. It is, in a sense, a choice to transform the relevant information into the symbolic domain for symbolic reasoning. The alternative is to transform the design and regulatory information into the geometric domain for geometric reasoning, which is the approach of Patlakas et al. (2018). Patlakas et al. (2018) checked the code and the model by presenting automatic code compliance of design calculation in building information models with a mathematical process, which allowed the substitution of a complex, multi-equation structural calculation algorithm with a single equation. Zhong et al. (2018) developed an ontology-based framework to check compliance with the building environment, geometry, and regulation data. Table 2-2 summarizes the literature in relation to automated compliance checking.

Table 2-2: Summary of the Literature Review on Compliance Checking

References	**Keywords**	**Strength**					**Weakness**				**Notes**
		F/M	**A/S**	**LR**	**DAP**	**BIM**	**LD**	**EK**	**T/D**	**NC**	
Yurchyshyna et al. (2009)	Ontology	X					X	X	X	X	
Yurchyshyna et al. (2010)	Ontology	X			X		X	X	X		
Eastman et al. (2009)	Geometric			X			N/A	N/A	N/A	N/A	Software included
Salama & El-Gohary (2011)	Regulation	X	X				X		X	X	
Salama & El-Gohary (2013)	Hierarchies	X			X			X	X	X	
Salama & El-Gohary (2016)	Text classification	X	X		X		X		X	X	
Zhang & El-Gohary (2012)	Grammatical rules	X					X		X		
Zhang & El-Gohary (2013)	Improvement loop	X			X		X		X	X	Requires initial input

Zhang & El-Gohary (2015)	Syntax tag	X			X		X		X		
Zhang & El-Gohary (2016)	Information extraction	X			X		X				
Zhang & El-Gohary (2017a)	First order logic	X			X		X	X	X		
Zhang & El-Gohary (2017b)	Integration		X		X	X			X		Limited to quantita-tive clauses
Zhou & El-Gohary (2017)	Dependency theory		X		X			X	X	X	
Patlakas et al. (2018)	Parameters	X	X			X	X	X	X		Software runs until it finds best fit
Zhong et al. (2018)	Ontology	X	X			X	X	X	X		Sensor data allows real-time complia-nce checking

F/M: Framework/Model; A/S: Automation/Semi-automation; LR: Literature Review; DAP: Details Added to Previous studies; BIM: Building Information Modeling applied; LD: Large Data needed; EK: Expert Knowledge required; T/D: Technical/Data limitation; NC: Not for Commercial use

Most studies regarding ACC developed frameworks and/or models. Some of them focused mainly on (semi-) automation, while others utilized BIM to achieve their objectives. There is a literature review that mostly focused on geometric compliance and introduces software packages for the ACC purpose (Eastman et al., 2009). Another literature proposed an improvement by adding a loop in a framework (Zhang & El-Gohary, 2013). Recently, for geometric compliance, software packages were used to find the best fit once a parameter is defined (Patlakas et al., 2018). Ontology was a concept that has been studied constantly with technologies or toolsets (e.g., sensors) used for data collection, which allows real-time compliance checking (Zhong et al., 2018). However, most past studies required large data and expert knowledge. Because of limitations in technical knowledge and data, accuracy was not guaranteed for all ACC problems. There are also many exceptions in human languages. Additionally, the methods are not for commercial use.

These past studies can be categorized according to methodologies. The first category uses ontology, deontic theory, and natural language processing (NLP) (Salama and El-Gohary 2013; Zhang and El-Gohary 2016; Zhong et al., 2018). Due to the technical limitations and exceptions in human languages, these studies are limited to functional demonstration and proof-of-principle. The second utilizes BIM (Eastman et al. 2009; Palakas et al. 2018; Zhong et al., 2018). These past studies have resolved some ACC complications but are focused on conforming to quantitative values. However, adequate accuracy has not been reached for commercial use. Adequate accuracy is application-specific and is driven by the cost of false positives and negatives; that cost must be outweighed by the savings from automation, in order to make the approach useful in practice. Thus, a practical framework for companies is needed that overcomes technical limitations. Methodologies to approach the conformance measurement automatically, without sharing confidential data, are perceived as useful because the construction industry is still typically reluctant to share private data.

2.4 Conformance Checking in Process Mining

Another topic that must be addressed is process mining. Process mining is a field that discovers, monitors, controls, and improves processes by finding patterns of activities (Van der Aalst et al., 2010). The field is relatively new, but the toolsets can also be applied in the construction industry. More specifically, it accompanies the conformance checking subfield, a subfield related to the workflow conformance measurement presented in this study. By reviewing relevant literature, the where and how behind the utilization of process mining is revealed.

Processes are considered vital in the construction industry. Employees need to follow the procedures in a suitable manner and at the right time. Measuring process conformance and performance is also critical in order to make decisions regarding whether to keep or modify the processes. Nowadays, most companies desire to save records for future reference. This is especially true when a workflow is implemented on a project and an event log is created and saved automatically in the database. An event log is a collection of events recorded in sequential order, each of which refers to an activity (Mannhardt, 2018). With the event log, the implementation of a workflow can be analyzed.

Apart from current auditing procedures where only a small set of sample data is evaluated, process mining enables events evaluation for all events (Van der Aalst et al. 2010). Therefore, the results may be more accurate, and the analyses more reliable and logical. Essentially, an event log must include case IDs, activities, and timestamps (Suriadi et al., 2017). Optionally, other relevant features for events may be added, such as status, costs, people, or locations. Since the event log data includes timestamps, sequences of events are recorded and can be tracked. The ability of process mining toolsets to deal with big data has enabled process mining to be used with some success in industries, such as healthcare, banking, and manufacturing (Rojas et al., 2016; Werner, 2017).

It has been difficult to apply process mining techniques to the construction industry due to barriers, such as data accessibility, paper-based characteristics, and workflow complexity related to project-specific parameters. However, the reluctance of the employees to engage with new technologies is somewhat fair given that process mining is a relatively new field and its accuracy in the construction industry is still questionable. Nevertheless, as workflows are managed and enforced automatically through the workflow management system (WMS), there is high potential for the use of process mining toolsets. Since the process mining field is actively researched and the toolsets are improving, the accuracy of the commercial software packages is expected to increase.

In construction, many steps are uniquely created on a case-by-case basis, and processes are shared with multiple companies. Thus, it is difficult to find event logs or to discover representative workflows (Van Berlo & Natrop, 2015). However, there are recent studies that discuss process mining in construction. Van Schaijk and Berlo (2016) introduced process mining into the AECFM industry and provided some examples. Zhang & Ashuri (2018) developed a systematic approach to make good use of design log data to monitor and measure the productivity of the design process.

While the process mining field includes various subfields to analyze workflows, process discovery and conformance checking are two main subfields. Process discovery refers to developing a discovered workflow from an event log by identifying sequential patterns (Van der Aalst, 2017). A discovered workflow is defined in this study as a visual workflow that is derived from an event log. Conformance checking refers to comparing discovered workflows or event logs against benchmark workflows. The benchmark workflow, which represents the intent, purpose, or goal, is sometimes called the priori process model, target model, reference model, workflow, or de jure model (Van der Aalst et al., 2010). The discovered workflow that may change over time due to the implementation of the benchmark workflow is sometimes called the discovered model, observed model, or de facto model (Mariscal et al., 2010).

Several studies related to conformance checking are described. Conformance checking identifies two types of discrepancies. The first type is unfitting log behaviour, which refers to behaviour that is observed in the event log but does not exist in the benchmark workflow model. The second type is additional model behaviour, which refers to behaviour that exists in the benchmark workflow model but is never observed in the log. The identification of unfitting log behaviour has been approached using methods such as the token-based method and the footprint-based method.

The token-based method takes one case at a time as input and tracks activities with tokens (Weijters and Van der Aalst 2001). This method identifies two types of unfitting behaviours: missing tokens and remaining tokens (Rozinat and Van der Aalst 2008). Then, the numbers of consumed, produced, missing, and remaining tokens are taken into consideration and are entered into equations. The limitation of this method is that it may not identify the minimum number of errors that can explain unfitting log behaviour (Adriansyah et al. 2011; Mannhardt et al. 2016).

The footprint-based method identifies, for each case in the log, the closest corresponding route (or trace) by the benchmark workflow, and it computes the matching of the log and the benchmark workflow. It shows the points of deviation~~s~~ between the two routes, and outputs pairs of matching events. The~~se~~ matching pairs are counted; however, the number of matching events is often too large to be explored exhaustively.

The limitations of the two methods just discussed are that they only identify differences at the level of individual routes, rather than at the level of behavioural relations observed in the log. These ideas inspired more studies for model-to-model comparison (Armas-Cervantes et al., 2016) and log-to-model comparison (Van Beest et al., 2015). García-Bañuelos et al. (2018) checked the conformance between the event log and benchmark workflow, creating event structures for both the event log and benchmark workflow, in order to align the matching. Table 2-3 summarizes literature related to process conformance.

Table 2-3: Summary of the Literature Review on Process Conformance

Reference	Keywords	Strength					Weakness		
		C/F	DA	A/ST	CD	PM	M/T	MV	TL
Rozinat & Van der Aalst (2008)	Token-based	X	X	X		X	X		X
Van der Aalst et al. (2010)	Audit					X			X
Adriansyah et al. (2011)	Cost-based	X				X	X		X
Weijters & Van der Aalst (2011)	Concurrent events		X			X	X		X
Golzarpoor et al. (2016)	Process	X			X				X
Mannhardt et al. (2016)	Multi-perspective		X			X	X		X
Van Schaijik & Van Berlo (2016)	AECFM			X	X	X		X	X
Golzarpoor et al. (2017)	Interoperability	X			X		X	X	X
Van der Aalst (2017)	Spreadsheet		X	X		X	N/A	N/A	N/A
García-Bañuelos et al. (2018)	Natural language	X		X		X		X	X
Zhang & Ashuri (2018)	BIM collaboration level				X			X	X

C/F: Concept/Framework; DA: Details Added; A/ST: Automated/Semi-automated Tools; CD: Construction Domain; PM: Process Mining; M/T: Mathematic/Theoretical; MV: Missing Validations; TL: Technical Limitations

Processes in the construction domain are related to diagnosis and improvement of existing processes. Thus, only general guidelines are provided without considering every possible situation that the project is in. Although some studies focus on standardizing the processes (Golzarpoor et al., 2016), it is still in the beginning stages. Other fields, such as banking, manufacturing, or healthcare, focus on process improvement with process mining technologies. The construction domain is slowly approaching the process mining domain (Van Schaijik & Van Berlo, 2016); however, in many cases validation is missing and technical limitations exist. Since the process mining domain itself is relatively new, it is actively researched by many researchers to this day.

2.5 Correlation between Practice Conformance and Project Performance

Conformance, compliance, and conformity are used as synonyms. Debates regarding which of these synonymous words to use prompted the International Organization for Standardization (ISO) to respond with a relevant standard. ISO 9000:2015 suggests the use of conformity over conformance or compliance for quality management systems. However, this thesis will use the term conformance for several reasons, the first being that compared to the definition of compliance (i.e., doing what one is told to do), conformance, which refers to the choice to do something in a recognized way, is more inclusive. Conformance embraces compliance because conformance goes beyond abiding by the law or fulfilling regulations; it also suits the scope of pursuing non-compulsory practice guidelines. Moreover, the term conformance is more familiar to other fields, such as process mining. The context of using conformance in process mining is for conformance checking. These contexts align with the objective of this study, which is measurement by comparison against practice guidelines. This study includes literature related to both conformance and compliance.

It is found that management practice guidelines and performances are correlated in healthcare (O'Malley et al., 2004). Not only is the importance of conformance discussed, but conformance is enforced by software systems in automated ways. Blaser et al. (2007) developed a software engineering process to achieve an IT alignment to healthcare process requirements. Mosadeghrad (2013) discussed quality improvements while defining quality as "conformance to specifications, requirements, or standards." It is implied that the quality of healthcare services cannot be improved unless they conform to the specifications, requirements, or standards. The gap between guidelines and the performance of healthcare facilities is constantly recognized and improved (Lucas et al., 2013).

In manufacturing, the relationship between conformance and performance is modeled (Maani & Sluti, 1990). Through such modeling, it was found that few significant correlations exist between manufacturing improvement goals and action programs in companies with no formal or written strategy (Gertsen et al., 2003). The results indicate that companies with formal, written manufacturing strategies translate their manufacturing goals into action programs. Conformance to action programs is positively related to quality performance improvement. There are attempts to solve product quality/reliability and manufacturing conformance aspects (Laugen et al., 2005). Furthermore, standard conformance in manufacturing determines whether the interpretation of the standardized terms used by software applications is consistent with semantics given by the standards. Deshayes et al. (2007) proposed a general architecture to design ontologies for standards integration and conformance in manufacturing engineering. When it comes to software and automation, in order to be in conformance, terms and definitions must be aligned before processing.

In banking, service quality is measured through conformance (Roth & Jackson, 1995). Relevant evidence is gathered on staff attitudes by identifying what elements should be in conformance (Cowling & Newman, 1996). Reference architectures (RA) are reusable architectures for artifacts in a banking domain. They can serve as a basis for designing new architectures, but also as a means for quality control during system development. Quality control is performed by checking the conformance of systems in development to company-wide RA. If performed manually, reference architecture conformance checking is a time and resource-intensive process. Reference architectures are defined based on reusable rules, consisting of roles and constraints on roles and role relationships. Conformance checking can be performed semi-automatically and continuously by automating important steps like the extraction of the actual application architecture, the binding of RA roles to the elements of specific application architecture, and the evaluation of the RA rules for an application architecture (Buchgeher & Weinreich, 2013). Inspired by these previous studies, semi-automating the steps of conformance measurement in the construction industry is investigated.

In construction, measuring practice conformance is an emerging need that is particularly crucial for improving the management of capital projects (Golzarpoor, 2017). Past studies have validated the positive relationship between construction management best practice conformance and project performance (Rankin et al., 2008; Shan et al., 2011; Meng, 2012; Yun et al., 2016). Studies have measured the relationship between combinations of cross-correlated or interdependent management practices and project performances (Nasir et al., 2012). Such management practices include material management, project planning, safety and health, IT automation, and first-level construction supervisory skills.

2.6 Knowledge Gap

While capital projects involve the documentation of practice guidelines in various ways, practical frameworks and toolsets that enable the measurement of conformance in a capital project audit are lacking. Process automation in conformance measurement is considered important in various domains, such as healthcare, manufacturing, and banking. Past literature indicates that capital project practice guidelines are abundant, and conformance and performance are positively related. However, challenges remain that must be overcome. First, self-recorded data is likely to entail human errors. The results can also be manipulated. Second, conformance to the practice guidelines is measured manually, which entails subjectivity. The results can be biased intentionally or unintentionally. Alternative studies that pursue automation also exist. While functional demonstrations and proof-of-principles validate the high accuracy of these alternative studies, the studies are not at a level of commercialization because real-world problems are more complex and have more exceptions. These knowledge gaps are addressed throughout this thesis.

Project performance improves when processes and practices are well-defined and measured. Judicious automation with current technology can support more efficient, accurate, and quick measurements. The representation of the knowledge gap is illustrated in Figure 2-1.

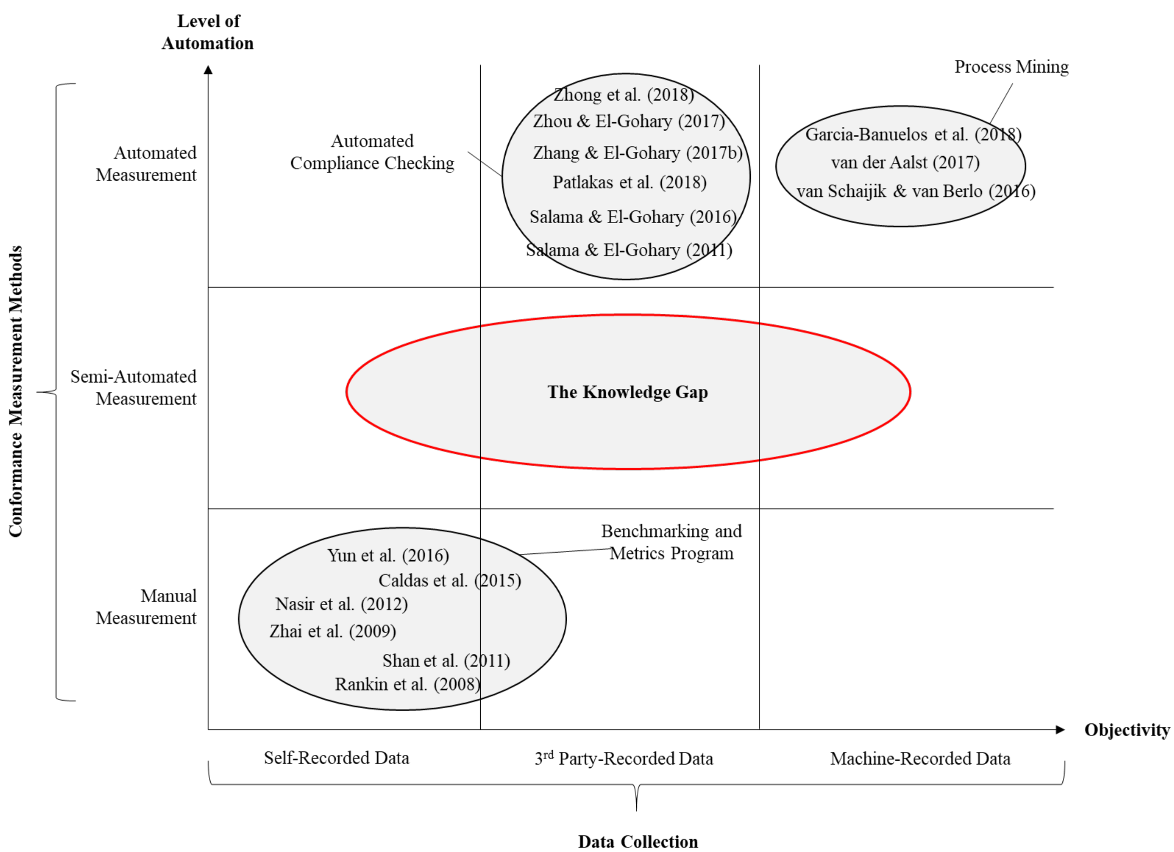


Figure 2-1: The Knowledge Gap

Literature related to practice conformance measurement is grouped into three categories: Benchmarking and Metrics Program (BM&M), Automated Compliance Checking (ACC), and Process Mining. The graph breaks down the literature in terms of objectivity on the x-axis and level of automation on the y-axis. A benchmarking and metrics program fits into self-reported data collection with manual measurement because the results are mostly derived from surveys. Surveys are simple and easy but human-errors are involved and can easily be biased. Automated Compliance Checking fits into 3rd party data collection. There are challenges in Automated Compliance Checking approach because of exceptions and anomalies for natural language problems. In process mining, data are collected from machine. Automated compliance checking and process mining target full-automation, but have yet to demonstrate their proficiency in automation within the practice, because they are in early stages. Process mining has not been applied to the construction industry because Return on Investment (ROI) has not been determined. Therefore, the knowledge gap is utilizing a combination of these methods.

Literature regarding benchmarking and metrics (BM&M) programs in the construction sector offers guidelines for how to approach the measurement of practice conformance and performance. However, BM&M programs face a dilemma when it comes to data collection, because such data is collected through in-person or telephone interviews with a corporate function or project/construction managers, a researcher's review of project files, and online or manual self-completed forms of industry participants. Data collection has difficulties due to time constraints and missing data, which is caused by concerns about revealing a competitive advantage. Moreover, the cost and time per unit do not always provide a meaningful comparison across projects, because such projects vary widely in nature. Differing definitions (e.g., safety incidents) and reluctance related to candid critiques (e.g., client satisfaction) have complicated thorough conformance investigations.

Currently, conformance measurement in the construction industry is mostly fulfilled by successfully conforming to compulsory guidelines, such as laws, regulations, or codes. Non-compulsory guidelines, such as best practices or benchmarks, are often not a priority. Some parts of Automated Compliance Checking (ACC) are achieved by the presence of Building Information Modeling (BIM) and Artificial

Intelligence (AI). However, these methods are still in the early stages. Therefore, practical and easier ways to meet the needs of construction practitioners are perceived to be crucial.

Another way to measure practice conformance is to evaluate every possible step of the workflow manually. However, considering that there are multiple processes in a project and limited man-hours, it is almost impossible for the auditors to track all events and documents in an entire process. To resolve this issue, process mining concepts and toolsets can be utilized; however, more validation is required because it is still at a proof-of-concept phase. Since it has seldomly been applied in the construction domain, challenges remain.

This thesis addresses the gap between: (1) research to-date that is focused on complete automation, which exhibits mostly inadequate accuracy, and (2) the need, as well as the opportunity, for automation with adequate levels of accuracy in specific practice conformance measurement activities. Thus, this thesis emphasizes semi-automation in construction management practice conformance measurement by presenting and testing potentially practical, reasonable, and consistent semi-automation solutions for measuring practice conformances. Such solutions can be a steppingstone for further exploration and can contribute to the body of knowledge for capital projects.

Chapter 3 A Practice Conformance Model

A model is a simplified representation of a system aimed to support an understanding of the real system. The process of establishing a model improves understanding of the problems and surrounding circumstances. It also identifies interrelationships between defined components of the model. To achieve practice conformance measurement, the development of a practice conformance model is perceived to be necessary. The practice conformance model presented here is developed from the analysis and synthesis of past studies, and the data are collected based upon the grounded theory research methodology. Rather than choosing between observing facts to form a model, and having a model to observe facts, an iterative approach is followed in this research. To make sense of the results, a model is presented first in the form of a capital project practice conformance model in this thesis. Detailed measurable elements are described after the illustrations and explanations of the practice conformance model.

3.1 General Components of the Practice Conformance Model

The practice conformance model is the foundation for the practice conformance measurement and the automation of the measurement process. In this thesis, it is asserted based on the literature review and inductive analysis of the facts that, documents, workflows, people, and actions are the minimum components that must be considered when defining and measuring practice conformance (Kang & Haas, 2018). Figure 3-1 illustrates practice conformance components, a summary of what can be measured, and the possibilities of automation in the near term.

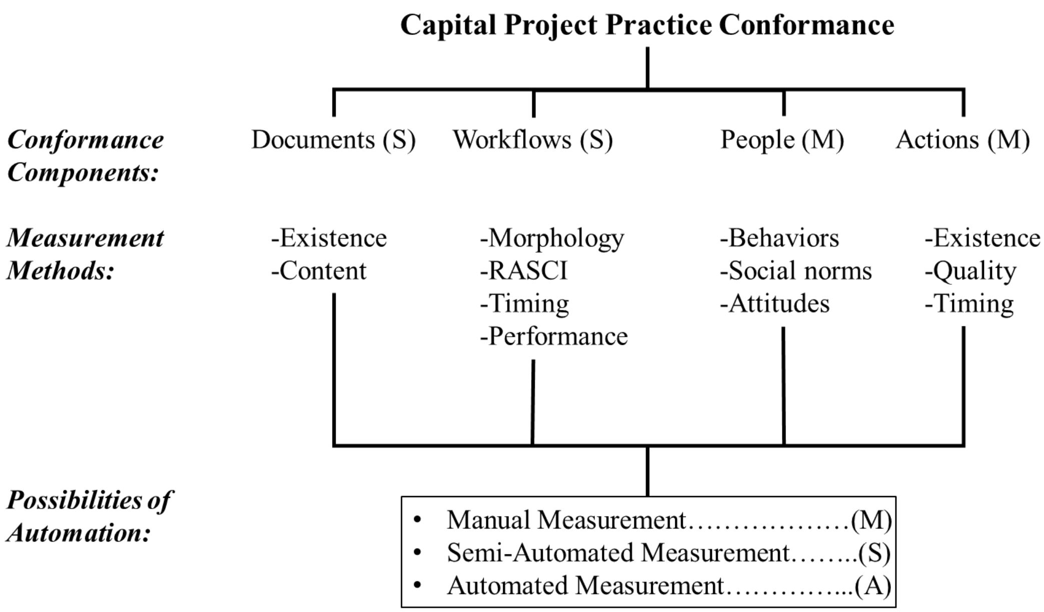


Figure 3-1: The Capital Project Practice Conformance Model (Initial)

1. Documents are paper or electronic objects. Examples of documents include contracts, specifications, drawings, 3D design files, and checklists. Documents may be annotated, have versions, and be associated with other documents, processes, projects, and people.

2. Workflows are outlines of processes that focus on the flow of organizational, operational, and implementation-level details (Golzarpoor, 2017). The flow of activities from the workflow can be updated.

3. People are defined from the perspective of their behaviour and attitudes, both individually and as groups. For example, a group might exhibit social norms that discourage the wearing of personal protective equipment in some situations, and that would be considered in this thesis as a form of measurement for lack of conformance to best safety practices.

4. Actions of both people and machines are defined here as measurable from a practice conformance perspective. For example, project participants may not be using job hazard

analysis, when the situation merits their use; so this does not conform to the best risk management and safety practices. As another example, a piece of heavy equipment may not sound an alarm when backing up, and that would be considered a non-conforming action (even if due to a malfunction).

Documents have more passive characteristics compared to workflows, which actively enforce the implementation of processes. Documents are essentially states, and workflows are the control logic for the transformations between states. Concerning people and actions, actions towards practice conformance may differ according to the mental state people are in. These practice conformance components are not exclusive but rather related and influence each other.

In Figure 3-1, the possibilities of automation in the near term are recorded (M: manual measurement, S: semi-automated measurement, A: automated measurement). Since documents and workflows are physical resources, semi-automation in near term is perceived as possible. Because there is not enough research and metrics to measure the practice conformance of people and actions, and because of the difficulties initiating the categorization of the components of people and actions, possibilities of automation for these components are graded low. Alternatively, the practice conformance of people and actions may be checked manually in the short-term. The reasons for the semi-automation of document and workflow conformance measurement are further explained.

Documents need to exist and be correct. For example, missing a part of contract documents is a red flag due to non-existence. In most capital projects, by using document management systems (DMS), documents are organized and managed. Sometimes the system does not allow a user to proceed unless all the files are attached. Therefore, auditors may easily check the existence of documents electronically. However, considering that not all documents are in electronic file forms in the construction industry, the auditors must acknowledge the existence of relevant documents.

Another issue is the existence within documents. Blank signatures, sections, and records are problems. For example, for project management documents, there must be authorizations of the documents either as a form of signature or stamp of a project manager. Checkboxes from a checklist must be marked. The existence of some elements within the document can be detected electronically. The current information technology (IT) systems prevent the users to proceed unless all signatures and checkboxes are filled out.

Likewise, the construction industry is pursuing automation. However, the need for automation in practice conformance measurement is not widely recognized yet. Other than determining the existence of specific elements of documents, detecting the correctness of the elements is more complicated and difficult. For example, to find out whether the signatures are received by the correct personnel and the checkboxes are checked in a considerate manner, artificial intelligence (AI) needs to be used, and big data is normally required to train the AI. Considering the current state of the construction industry, where big data can hardly be collected due to the reluctance of sharing the data, complete automation of measuring correctness remains a challenge.

A workflow typically includes morphology (core structure, abstraction level), RASCI (roles and responsibilities: Responsible, Accountable, Support, Consulted, Informed), and timing (expected durations). Workflows may be illustrated manually or electronically. Once the workflow is designed, it must be implemented and timing (actual durations) and performances need to be measured. It is essential to define key performance indicators (KPIs), based upon the goals and objectives of a company, prior to measurement.

When the workflow is created manually as a flowchart, it is more difficult to measure the practice conformance. When the workflow is enforced by the system electronically, it is easier to control the process because some activities are prohibited automatically by the system. Hence, practice conformance is reached. For example, assume that a management process is achieved electronically. When the management benchmark workflow is created, the morphology or the collection of relations between activities is developed and saved. Then, roles and responsibilities are designated to relevant parties, and they are saved to the management system. In this way, some participants have the rights to modify whereas others only have the right to view attachments. Practice conformance is already achieved partially by applying the electronic system.

Practice conformance measurement is also easier when performed electronically because the system can track which participant has completed which activity. However, due to the complexities of, and exceptions that exist in workflows, practice conformance to the benchmark workflow is still difficult to measure.

Defining practice conformance and developing appropriate metrics for practice conformance measurement are some tasks that must be achieved. The performance is also recommended to be defined and measured. Typically, in capital projects, higher performance signifies higher quality, shorter time, and/or lower costs. When it comes to workflow conformance, quality is perceived harder to measure than durations and costs. This is because oftentimes subjectivity is involved, and definitions may be vague for higher quality.

Additionally, there may be no direct data on costs or timing. In these cases, the costs can be defined as the number of personnel involved to complete a process. Timing can also be a measure of performance. Three types of timing may be measured: activity time, which is the duration from the activity start to the activity end; waiting time, which is transition duration from one activity end to the next activity start; and deadline, which is the total duration from the start of the first activity to the end of the last activity. These actual durations can be compared to the planned (expected) project timeline.

Documents and workflows are created, executed, and completed by people. People are not final deliverables or outcomes. However, they take part in every process. They influence document and workflow conformances. Their behaviours, social norms, and attitudes are the elements that define an individual or a society, and these elements allow one to forecast actions to some extent. For instance, when a construction site is unclean and messy, workers and managers may be the source due to their unwise habits, lack of social norms, or passive personalities. These may have caused the non-conformant actions. By including people as part of the practice conformance model (Figure 3-1), the causes of non-conformance can further be revealed. Investigating behaviours, social norms, and attitudes of people may be ways to predict and improve practice conformances. However, the people component is perceived difficult to measure especially in automated ways. The first step is to create a list that is specific and repeatable.

Lastly, when measuring an action component, its existence, correctness (quality), and efficiency (timing) must be measured. Actions are taken by people and machines. However, there are two reasons for this distinction between people and actions. The first reason is that people do not necessarily perform their actions in certain ways. Peoples' behaviours, social norms, and attitudes certainly affect actions. However, their actions can also change according to company policies or other factors. The second reason is that actions are what actually happened as opposed to what attitudes people have. Actions are affected by people and influence the outcome of documents and workflows. The following questions may be asked to measure actions: "Do the actions exist?"; "Are these actions the correct actions?"; and "Are these actions efficient enough?"

Manual measurement (M) refers to the current prevalent audit approach, wherein several professionals go through all the documents and workflows, and/or investigate people and actions (Figure 3-1). Typically, only limited samples are collected due to time constraints. Semi-automated measurement (S), for which this research aims and provides framework, allows human involvement. Some auditing tasks related to practice conformance measurement are simple but tedious and time-consuming. These tasks must be completed typically before actual analyses. On the other hand, some tasks require logical thinking and judgment. Sampling and analyzing are examples of these tasks. If tasks are divided into ones that are easy to measure manually and others that are more challenging whether and how to automate these tasks can be decided. For example, tasks that require senses (e.g., sight, hearing, smell, and touch) are easier tasks that can be manually measured. However, because of the large amount of information and complexity, the assistance of automation can be significantly useful. Hence, tasks that require senses may be attempted to be automated in the near term, but, details are explained later in the chapter. Full automation (A) of analyzing and decision making may take longer to accomplish. Though core technologies are growing and improving, replacing the audit requires much data (for machine learning, for example) and higher accuracy.

This thesis aims to establish a framework for measuring practice conformance for capital projects in semi-automated ways. A framework and toolsets to assist practice conformance measurement are in high demand. To attain the precise results of practice conformance, the practice conformance model must also be structurally developed and validated.

3.2 Capital Project Practice Conformance Elements

Capital project practice conformance elements are developed to specify the practice conformance model (Figure 3-1). Capital project practice conformance elements are defined as measurable indices of practice conformance components (i.e., documents, workflows, people, and actions). Practice conformance elements are derived based on the grounded theory research methodology. This research methodology takes a normative approach whereby value-based, and qualitative data creates a theory (Glaser & Strauss, 1967). By applying this methodology, the practice conformance model is strengthened with concrete measurable elements and definitions. The process, amount, and quality of data collection are critical for this approach.

3.2.1 Derivation of the Practice Conformance Elements

To understand construction management practice conformance from an industry point of view, interviews with several professionals from the industry (Total: 5) were conducted based on the grounded theory research methodology. The grounded theory method was first proposed by Glaser and Strauss in 1967. They defined the grounded theory as a discovery of theory from data that is systematically obtained and analyzed in social research. By applying this concept, the practice conformance model is fortified. To gain broad opinions and ideas, unstructured interviews were conducted. A professional refers to a person with competency and skills in a particular industry sector. A total of five professionals within the industry can find 85-90% of usability problems. Utilizing more than five can be redundant which is why five professionals were interviewed in this study. This representation of findings is adequate according to Nielson and Landaur (1993).

Built upon the practice conformance model, qualitative data from interviewees was collected. Then, researchers reviewed and reorganized data. As the interviews went on, repeated ideas, concepts, or elements became apparent. Data were then re-reviewed and grouped into concepts and categories. These categories are the foundation of practice conformance elements. Practice conformance elements are derived based on industry point of view, apart from the practice conformance components that are derived through inductive analysis. This approach from both perspectives reduces gaps between academia and industry. Table 3-1 presents the results of consultations and meetings to identify the practice conformance elements. A (+) sign indicates that the element is included in the list, (*): modified, and (-): removed.

Table 3-1: Derivation of Practice Conformance Elements

Date	**Elements** (+) : Added; (*) : Modified; (-) : Removed	**Participant Information** (The Number of People)	**Reasoning / Comments**
2017. 07. 18	(+) Core Structure (+) Abstraction Level (+) Participants (+) Time (+) Data	Internal meeting (3) (Professor, Post Doc, PhD)	• Literature review
2017. 08. 10	(+) Keywords (*) Numbers (+) Examples (*) Processes	Internal meeting (3) (Professor, Post Doc, PhD)	• **Time** can be generalized to **Numbers** • **Processes** include **Core Structure** and **Abstraction Level**
2017. 08. 10	(+) Forms/Checklists	Expert interview (2) (Contractor)	• **Forms/Checklists** make it clear what should be done
2017.08. 14	(*) Workflow Morphology (*) Words in Descriptions (*) RASCI (+) Human Activities/ Implementation	Internal meeting (3) (Professor, Post Doc, PhD)	• **Processes → Workflow Morphology** • **Keywords → Words in Descriptions** • **Participants → RASCI** • Not only the process but implementation should be considered
2017. 08. 15	(+) Links	Expert interview (2) (Owner)	• Most documents refer to other documents
2017. 08. 16	(*) Description/Narratives (*) Roles/Responsibilities (*) Processes/Flowcharts	Expert interview (2) (Consultant)	• **Words in Descriptions → Description/Narratives.** • **RASCI → Roles/ Responsibilities** • **Workflow Morphology →Processes/Flowcharts**
2017. 08. 18	(*) Table of Contents	Expert interview (2) (Supplier)	• **Table of Contents** mostly exists and is the first thing to look at
2017. 08. 25	(+) Bullet Points (+) Figure/Photo (+) Chart	Internal meeting (3) (Professor, Post Doc, PhD)	• *Added* elements that may exist in documents
2017. 08. 28	(-) Human Activities/ Implementation (-) Examples (-) Data	Internal meeting (3) (Professor, Post Doc, PhD)	• *Removed* elements that are vague or not measurable
2017. 09. 14	(*) Process/Flowchart (*) Roles of Participants (*) Deadlines/Durations (*) Figures/Photos (*) Charts/Tables (*) References	External meeting (3) (Owners, PhD)	• *Modified* terms to facilitate understanding • **Roles/Responsibilities → Roles of Participants** • **Time → Deadlines / Durations** • **Links → References**
2017. 09. 22	(+) Revision Number/Date (+) Document Approval / Authorization	External meeting (3) (Owners, PhD)	• *Added* elements that are measurable • *Modified* terms to facilitate understanding

	(*) Itemized List of Activities (*) Benchmark Historical data (-) References		• *Deleted* element that overlapped meaning with other elements • **Bullet Points → Itemized List of Activities** • **Numbers → Benchmark Historical Data**
2017. 10. 31	(*) Process Flowchart (+) Definition/Acronym (*) Relevant Organization Chart/RASCI (*) Schedule/Timeline (*) Relevant Project Form/Checklist (*) Relevant Chart/Graph (*) Relevant Photo/Image	<u>External meeting (4)</u> (Owner, Contractor, Professor, PhD)	• *Added* elements that are measurable • *Modified* terms to facilitate understanding

This process of adding, removing, and modifying elements is to extract valuable and measurable elements and accurately deliver information to the industry practitioners using their language. By having input from industry professionals with interviews and meetings, the list is refined.

In the early stage, elements of process conformance were identified prior to practice conformance (Golzarpoor et al., 2016a). The "core structure" and "abstraction level" are the main elements to be aligned for the processes to conform to. When these two elements align, "participants," "time," and "data" that are transferred are perceived to be checked in detail quantitatively and qualitatively. However, after a few meetings, it was decided to embrace practice conformance. Therefore, elements that consist of capital project practices were identified and discussed.

Simple terms such as "keywords" and "examples" imply the conformance of semantics of documents. When they are discussed, it is suggested to first check whether these elements exist in the documents. If "examples" exist in documents, it means the document has been customized according to the company situation, which is a positive sign. If the "keywords" of corporate documents match with the practice guideline documents, it is also a positive sign in terms of alignment. However, it is recommended that auditors be cautious of the manipulation of documents. For example, auditees may expect to achieve conformance by adding a few "keywords", regardless of appropriateness. This action must be identified and sanctioned. However, there also need to be warnings when corporate documents are highly creative. Hence, conformance must be defined before performing the audit.

To have a better understanding of practice conformance elements, more meetings occurred. “Numbers” imply quantitative comparison, and “processes” are an aggregation of the “core structure” and “abstraction level.” These are all related to the practice conformance components from the practice conformance model. “Processes” element is modified into “workflow morphology.” Workflow is described as a visual outline of a process (Golzarpoor 2017), and morphology is referred to as a form of the core structure. This changes into “processes/flowcharts” then to “process flowcharts.”

“Forms and checklists” is an important interactive element because it requires responses. These are related to people and actions components. According to the responses, implementation of practices can be measured. Whereas the existence of checkmarks on a checklist can be measured, the quality of responses cannot be objectively measured if they are self-reported. When the third-party observes and marks the checklist, the risks arising from self-reporting can be reduced. Inspired by “forms and checklists,” “human activities/implementation” was also discussed and was suggested to be included. However, in the end, the “human activities/implementation” element was excluded from the practice conformance elements because it is perceived to be unmeasurable.

“Keywords” and “examples” were modified into “words in descriptions” and then into “description/narratives.” “Participants” were modified and classified into “RASCI.” “RASCI” is an acronym for Responsible, Accountable, Supportive, Consulted, and Informed. A RASCI chart typically shows the roles and responsibilities of participants. Therefore, “RASCI” was modified into “roles and responsibilities.” In the end, this term changed again into “organization chart/RASCI.”

Some documents are created from the computer and online sources and are referenced often. Even within the companies, lots of links are associated with a document because it is more convenient to link the related topic than create a whole document on the same topic. “Links” was later modified to “references”, which was deleted in the end because it has an overlapped meaning with the “table of contents” and the “itemized list of activities.” “Table of contents” element is included as they typically exist at the front of documents indicating what is included in the document and what is not. More elements such as “bullet points,” “figure/photo,” and “charts” were discussed since some of these elements are included in practice conformance elements. These elements were modified as the “itemized list of activities,” “relevant photo/image,” and “relevant chart/graph.”

The “time” element changed into “deadlines/durations” and then into “schedule/timeline.” In the last few meetings, elements such as “revision number/date” and “document approval/authorization” were added for they are most often included in practice documents and added validities. “Numbers” were modified into “benchmark historical data” for clarification. In practice documents, companies are required to record numeric values to specify their goals or results in some cases. These numeric values are referred to as the “benchmark historical data.”

Several terms of elements were modified and others were removed from discussed elements. However, after twelve meetings, the final capital project practice conformance elements were concluded. The final practice conformance elements, their definitions, and potential uses for measurement are presented in Table 3-2.

Table 3-2: Final Practice Conformance Elements, Definitions, and Potential Use for Measurement

No.	Capital Project Practice Conformance Element	Definition	Potential Use for Measurement
1	Schedule/Timeline	A project/company plan for carrying out a process or procedure, giving lists of intended events and times	If it exists, it defines the timing and dependencies, clear resource identification, and event outline.
2	Process Flowchart	Well-defined steps of a process in sequential order	If it exists as a drawing in a document, it demonstrates institutional intent; if it exists as the markup for an automated business process workflow engine, it demonstrates intent and some elements of enforcement.
3	Relevant Organization Chart (RASCI)	A relevant graphic representation of the structure of an organization, showing the relationships of the positions or jobs within it (Abbreviation for Responsible, Accountable, Supportive, Consulted, and Informed)	If it exists, it represents a clear definition of all required positions and individuals' awareness of their responsibilities.
4	Itemized List of Activities	Itemized documentation of all of the activities that are part of a project	Existence and correctness allow to quickly understand the scope at a high level.
5	Description/Narratives	Written account of connected events	Existence and correctness in documents demonstrate intent, purpose, and scope of work.
6	Document Approval/ Authorization	Written consent by an authorized party	Existence establishes commitment and authority.
7	Table of Contents	A list, usually found on a page before the start of a written work, of its chapter or section titles or brief descriptions	Existence of such list ensures the flow at a high level.
8	Relevant Project Form/ Checklist	A relevant list of items required, things to be done, or points to be considered, used as a reminder	Existence and correctness prove a degree of practice conformance exists, ensuring critical steps are not missed.
9	Benchmark Historical Data	Past-periods data used usually as a basis for forecasting the future data or trends	Existence of such data demonstrates a degree of practice conformance exists.
10	Revision Number/Date	A version number/date when changes are made	Non-existence implies either no revision or no standard for indicating versions, a form of non-conformance.
11	Definition/Acronym	A statement of the meaning of a term/An abbreviation formed from the initial letters of other words and pronounced as a word	The correctness of such data satisfies practice conformance establishing a common understanding by aligning terms.

12	Relevant Chart/Graph	A relevant sheet of information in the form of a table or diagram	Existence typically provides additional information.
13	Relevant Photo/Image	A relevant graphic representation of the external form of a person, place, or thing	Existence typically provides additional visual information.

Table 3-2 explains the definition of each capital project practice conformance element and how it can be used as measurable indicators. The existence of an element is a critical dimension because it implies that the information exist within the element (or it can be inferred that the information exists due to the existence of the element). For example, the existence of a "schedule/timeline" validates that the company has a detailed plan. It implies that the tasks have timeframes, shows the dependencies of the tasks, identifies resources, and outlines activities.

A "process flowchart" can exist in two ways. If it exists as a drawing (i.e., flow diagram) in a document, it mostly demonstrates the intent of a company, in the form of a high-level practice, which leaves room for judgment of practitioners. If it exists as an automated business process workflow engine, it not only demonstrates the intent of the company but also includes a specific step-by-step process enforced to some extent by the logic of the workflow engine and the IT system in which it operates. That is, the next step cannot be achieved unless the previous steps have been executed properly.

When a "relevant organization chart/RASCI" is included in the practice, the positions and responsibilities of employees are defined at the same time. When "document approval/authorization" and "revision number/date" exist, it is already close to achieving the practice conformance. For example, approval of a project change order accepted by access-authorized members of the organization implies practice conformance since there are responsible individuals or departments. Apart from correctness, these examples reinforce the notion that the existence of some elements simply increases practice conformance. In these cases, practice conformance does not necessarily require a best practice, because there is no right or wrong answer. Simple existence is meaningful, for it allows reasonable measurement.

“Relevant project form/checklist” is an interactive element that the respondents use (sometimes collaboratively). Therefore, it may include additional information, such as opinions, perceptions, implementations, and performances. These checkmarks or responses from respondents may not be objective or candid; however, filling out the blanks (i.e., existence) is one form of practice conformance, in that the participants are aware of the element.

When it comes to correctness, there must be best practices or at least something that actual performance can be compared to. For example, comparisons of the “table of contents” may identify the differences and recognize the missing or additional information from best practices. Sometimes disclosures, statements of warranty or disclaimers must be present even if they are identical to general ones. Construction safety guidelines must include a clause such as “employees must wear personal protective equipment” or a chart for drug or alcohol thresholds. Missing these elements raise a red flag. Identical terms, phrases, or images are sometimes necessary. The correctness (i.e., the quality of the elements) is more difficult to measure than existence; therefore, it requires experts or sophisticated tools. The correctness of two identical elements is more amenable to measurement than potentially similar elements.

There are many issues when identifying synonyms, similar images within contexts, or customized workflows. Some may argue that the documents, workflows, and actions are in conformance while others may argue they are not. Because of this subjectivity, practice conformance must be well-defined so that it can be measured in quantitative and repeatable ways. Logical thinking processes and statistical analysis fall under the artificial intelligence (AI) category. This is still an ongoing research field with sub-categories, such as natural language processing (NLP) and image processing. Application to specific domains such as construction is currently limited. More importantly, big data is necessary for AI training. Apart from the construction industry, knowledge and information are shared rapidly through the internet. However, even then, the key strengths of companies are kept confidential with a non-disclosure agreement. When it comes to the construction industry, companies are more reluctant to share their knowledge or information. In later chapters, how this thesis avoids such issues is explained.

3.2.2 Validation of the Practice Conformance Elements

To validate the conformance elements, a survey was conducted after interviews and meetings. Based on the grounded theory research methodology, data was collected systematically for analysis. The results validated that the measurable conformance elements exist. The results also implied that the automation of certain conformance elements would have greater impacts than others. The survey provided useful information on how individual employees within certain categories view the conformance elements and how these categories can be considered as a generalized perspective.

3.2.2.1 Demographics of Survey Participants

From August, 2017, to November, 2017, a survey was designed to estimate the relative impacts of the practice conformance elements. Impact is defined as the degree to which an element may indicate conformance to a practice guideline. The intent was not only to validate the practice conformance model but also to support the framework for automation. This survey successfully confirms that there are measurable practice conformance elements and that the needs to measure these conformance elements vary. Elements are assessed in terms of the level of importance, frequency of use, and degree of effort to understand.

The process of the Office of Research Ethics (ORE) at the University of Waterloo was followed to construct the survey questionnaire. The ORE process requires submitting the ORE101 form (ORE#22565) as well as the information letter, consent form, and feedback letter. This process supports confidentiality as well as the formality and appropriateness of this survey. Table 3-3 summarizes eight demographic questions asked of the survey participants, and it outlines options and reasons behind them.

Table 3-3: Summary of the Participant Demographic Questions

Question Type	Options	Rationale
Industry Sector (Industry sector in which the participant has the most experience)	• Upstream & Midstream (Oil and Gas) • Downstream & Chemicals (Refining, Petrochemical, etc.) • Mining • Power, Utilities, and Infrastructure • Institutional/Commercial Facilities • Manufacturing	• Adopted classification of a recognized organization (i.e., Construction Industry Institute). • However, some differences exist. Mining is separated from Upstream & Midstream (Oil and Gas) based on end products. Institutional Facilities and Commercial Facilities are merged since their construction processes are alike.
Stakeholder Group (Stakeholder group in which the participant has the most experience)	• Owner • EP/EPC/EPCM • Contractor • Supplier • Public Sector	• Options are based on roles and purposes. • EP (Engineering & Procurement), EPC (Engineering & Procurement & Construction), and EPCM (Engineering & Procurement & Construction & Management) are bundled as their roles are alike. • Contractor is separated from EP/EPC/EPCM, and Public Sector is separated from Owner since the main purposes differ.
Project Phase (Project phase in which the participant has the most experience)	• Feasibility and Conceptual Planning • Design and Engineering • Construction • Commissioning • Operation and Maintenance	• Adopted classification of a recognized organization. • However, some differences exist. Procurement phase is removed as it is associated with all phases.
Role (Role in which the participant has the most experience)	• Project Management • Construction Management • Engineering Management • Procurement/Supply Chain Management • HSE Management • Material Management • Project Controls • Quality Management	• Derived from the consultations and meetings with experts. • HSE (Health, Safety, and Environment). • The main role of Project Control is to estimate costs and schedules.
Work Location (Work location in which the participant has the most experience)	• Site/Field-based • Corporate/Home Office-based	• Derived from the consultations and meetings with experts.

Participant's Age (optional)	<30 30-39 40-49 50-59 >60	• Divided into five options with ten-year gaps to aim at a good distribution.
Participant's Experience in the Industry	<5 years 5-10 years 11-20 years 21-30 years 31-40 years >40 years	• Divided into six options with five- or ten-year gaps to aim at a good distribution.
Participant's Level in Organization	• Individual Contributor • Mid-level Management • Senior-level Management	• Derived from the consultations and meetings with experts.

These questions are significant, since diverse participant categories may have unique needs and desires. The survey was launched on January 2nd, 2018, and the survey data was collected until January 31st, 2018. In total, sixty-one participants responded from distribution to 143 people resulting in a 42.7% response ratio. Forty-two responses were received electronically, and nineteen responses were collected in paper format. The distribution was assisted by the members of Construction Owners Association of Alberta (COAA) and Construction Industry Institute (CII). The demographic information of participants is shown in Table 3-4. These general characteristics are derived with rationales similar to those used in previous related studies and consultations (Table 3-3).

Table 3-4: Demographics of Survey Participants

Category		**Respondents**	
		Number	**%**
Industry Sector	Upstream and Midstream (Oil and Gas)	32	52
	Downstream and Chemicals (Refining, Petrochemical, etc.)	14	23
	Mining	1	2
	Power, Utilities and Infrastructure	3	5
	Institutional/Commercial Facilities	4	7
	Manufacturing	2	3
	Others	5	8
	SUM	**61**	**100**
Stakeholder Group	Owner	18	29
	EP/EPC/EPCM	9	15
	Contractor	30	49

	Supplier	2	3
	Public Sector	1	2
	Others	1	2
	SUM	**61**	**100**
Project Phase	Feasibility and Conceptual Planning	4	7
	Design and Engineering	11	18
	Construction	36	59
	Commissioning	2	3
	Operation and Maintenance	8	13
	SUM	**61**	**100**
Role	Project Management	17	28
	Construction Management	19	31
	Engineering Management	2	3
	HSE Management	2	3
	Material Management	2	3
	Project Controls	8	13
	Quality Management	4	7
	Others	7	12
	SUM	**61**	**100**
Work Location	Site/Field-based	29	48
	Corporate/Home Office-based	32	52
	SUM	**61**	**100**
Age	<30	14	23
	30-39	10	16
	40-49	15	25
	50-59	19	31
	>60	3	5
	SUM	**61**	**100**
Experience in Industry	<5years	11	18
	5-10 years	11	18
	11-20 years	14	23
	21-30 years	13	21
	31-40 years	10	17
	>40 years	2	3
	SUM	**61**	**100**
Level in Organization	Individual Contributor	15	24
	Mid-level Management	26	43
	Senior-level Management	20	33
	SUM	**61**	**100**

For the industry sector, more than half of the participants have most experiences in upstream & midstream, and downstream & chemicals. This may be because most respondents are from the United States (Texas region) or Canada (Alberta region). Since questionnaire distributions were assisted by the members of Construction Industry Institute (CII) and Construction Owners Association of Alberta (COAA) which are based in Texas and Alberta respectively, the distribution of respondents were expected. For the stakeholder group, more than 90% are categorized as owners, EP/EPC/EPCM, or contractors. More than 75% of participants have more experience in design & engineering or construction. More than half of the participants have more experience in project management or construction management. Their work locations, ages, experience years, and levels in organizations vary.

3.2.2.2 Validation of Practice Conformance Elements

In order to validate practice conformance elements derived by industry experts, three types of questions were asked for Part 2:

- The "practice conformance element" is important.
- The "practice conformance element" is consistently used in practices.
- The "practice conformance element" requires considerable efforts to understand.

Additionally, to reduce confusion and to crosscheck the responses gained from Part 2, Part 3 question is included:

- Considering the importance, consistent usage, and efforts needed, which practice conformance elements should gain priority over others? Please rank the number from 1 to 13 and provide reasons.

There are multiple intentions of asking three criteria (i.e., importance, usage, effort). The main intention is to investigate the opinions and perceptions of participants on practice conformance elements. There are also intentions to discard survey responses that have awkward patterns in their responses, such as all 5s or 12345 in sequence markings. These questions leveraged and filtered unintended responses. If any respondent responds with inconsistencies, the response was investigated carefully. Because the respondents answered with integrity, no survey response was discarded.

The importance of a practice conformance element refers to the level of significance perceived by participants when the practice conformance element is measured. The usage of a practice conformance element refers to the occurrence of the practice conformance element when the conformance is measured. The effort refers to the level of effort required by participants when the practice conformance elements are measured.

Depending on which practice conformance element is important, is consistently used, and is difficult for humans to understand, the potential impact of an element can be appraised. A practice conformance element with higher potential impact indicates that the element's conformance is more influential in overall conformance measurement than elements with lower potential. While it may be easier to start automating simple practice conformance elements and their measurement, the priorities of the automation are revealed by the survey. The results of this survey are the foundation of the capital project practice conformance model.

The analysis is computed by the use of the Statistical Package for the Social Sciences (SPSS). Relative Importance Index (RII) with five Likert scales (1: strongly disagree, 5: strongly agree) is adjusted in meaning here, because of the form of the questions. Instead, it is used as a Relative Weight Index (RWI). The Relative Importance Index (RII) is often used in construction-related research to investigate perceptions on factors affecting conditions or situations (e.g., delays (Gündüz et al., 2013)). Using Equation (1), the results are presented in Table 3-5, Table 3-6, and Table 3-7.

$$\text{RWI} = \frac{\sum W}{A \times N} \; (0 \leq RWI \leq 1) \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where:

- W: the weight given to each practice conformance element by the participants ranges from 1 to 5 (where 1 strongly disagrees and 5 strongly agrees);
- A: the highest weight (i.e., 5 in this case) and;
- N: the total number of participants

Table 3-5, Table 3-6, and Table 3-7 lists RWI in descending order to demonstrate which elements are considered more important, are perceived to be used more frequently, and are more difficult to understand than others. Then, the conformance elements are grouped into three levels of impact of automation.

Table 3-5: Survey Results – "Level of Importance" Order

Practice Conformance Elements	**Responses**					**RWI**
	Strongly Agrees	**Somewhat Agrees**	**Neither Agree nor Disagree**	**Somewhat Disagrees**	**Strongly Disagrees**	
	5	**4**	**3**	**2**	**1**	
Schedule/Timeline	49	10	2	0	0	0.95
Document Approval/ Authorization	44	11	4	2	0	0.92
Revision Number/Date	40	16	5	0	0	0.92
Itemized List of Activities	39	16	5	1	0	0.91
Relevant Organization Chart/RASCI	37	18	5	1	0	0.90
Benchmark Historical Data	37	19	2	2	1	0.89
Relevant Project Form/ Checklist	38	14	7	2	0	0.89
Table of Contents	38	14	6	3	0	0.89
Description/Narratives	37	16	5	2	1	0.88
Process Flowchart	29	27	4	0	1	0.87
Relevant Photo/Image	27	19	14	1	0	0.84

Definition/Acronym	26	21	7	3	4	0.80
Relevant Chart/Graph	13	26	16	6	0	0.75

For the importance criteria, "schedule/timeline" turns out as the most important element since no participant disagreed. "Document approval/authorization" and "revision number/date" follow in importance. However, almost all elements scored more than 0.8, implying that all the elements are more or less perceived to be important.

When it comes to the level of importance, the higher the rank is, the greater the impact is expected. Therefore, if the conformance measurement of the "schedule/timeline" is automated, it has a greater impact than the conformance measurement of the "relevant chart/graph." However, the level of importance is not the only consideration factor. Even if it is not important, if it is frequently used, the automation of conformance measurement of that element may be equally impactful. Therefore, the frequency of use of the practice conformance elements was measured (Table 3-6).

Table 3-6: Survey Results – "Frequency of Use" Order

Practice Conformance Elements	**Responses**					**RWI**
	Strongly Agrees	**Somewhat Agrees**	**Neither Agree nor Disagree**	**Somewhat Disagrees**	**Strongly Disagrees**	
	5	**4**	**3**	**2**	**1**	
Schedule/Timeline	33	21	4	3	0	0.88
Table of Contents	33	18	6	4	0	0.86
Revision Number/Date	32	18	8	3	0	0.86
Document Approval/ Authorization	27	26	7	1	0	0.86
Definition/Acronym	29	20	8	3	1	0.84
Itemized List of Activities	24	24	11	2	0	0.83
Relevant Organization Chart/RASCI	24	19	13	5	0	0.80
Description/Narratives	22	22	12	4	1	0.80
Relevant Photo/Image	13	30	10	6	2	0.75
Process Flowchart	13	25	17	5	1	0.74
Relevant Project Form/ Checklist	13	28	10	10	0	0.74
Benchmark Historical Data	12	19	12	12	6	0.66
Relevant Chart/Graph	2	28	18	13	0	0.66

When considering the “frequency of use” criteria, the first three elements from the “level of importance” criteria remain, adding the “table of contents” in the second rank. The results indicate that the “schedule/timeline” is frequently used, whereas the “relevant chart/graph” is not as frequently used. The matching results from Table 3-5 and Table 3-6 imply that the participants’ perceptions of the “level of importance” and the “frequency of use” are correlated.

Because the companies have limited time and resources, level of importance and frequency of use are not the only factors for automation. The difficulties of understanding the practice conformance elements must be considered as well. The last question is not phrased as the degree of difficulty of measuring the practice conformance elements but as the degree of difficulty of understanding the practice conformance elements (Table 3-7). This is intentional because the survey participants are not auditors. Rather than forcing the participants to assume they are auditing, explaining their own experiences was perceived more useful.

Table 3-7: Survey Results – “Degree of Effort to Understand” Order

Practice Conformance Elements	**Responses**					**RWI**
	Strongly Agrees	**Somewhat Agrees**	**Neither Agree nor Disagree**	**Somewhat Disagrees**	**Strongly Disagrees**	
	5	**4**	**3**	**2**	**1**	
Relevant Chart/Graph	9	31	13	7	1	0.73
Benchmark Historical Data	12	23	9	15	2	0.69
Schedule/Timeline	11	20	10	15	5	0.66
Definition/Acronym	7	17	16	9	12	0.59
Description/Narratives	4	18	14	19	6	0.58
Process Flowchart	5	20	7	21	8	0.58
Relevant Project Form/ Checklist	5	11	16	25	4	0.56
Relevant Organization Chart/RASCI	7	14	6	20	14	0.53
Itemized List of Activities	7	8	14	22	10	0.53
Table of Contents	7	4	7	17	26	0.43
Revision Number/Date	6	3	8	19	25	0.42
Document Approval/ Authorization	4	4	9	19	25	0.41

Relevant Photo/Image	2	2	13	12	32	0.37

When humans have difficulties understanding certain elements, there are two possibilities. Computer algorithms are better or worse than humans at the comprehension of an element. It is the same with automating the conformance measurement of practice conformance elements. It is hard to say which element will be more effective or more efficient. However, it is worth understanding the stance of professionals as a starting point to assess automation potential. Table 3-7 provides information for making decisions about which practice conformance elements to automate. When automation of one element is easier or quicker to achieve than others, it may be the priority. On the other hand, if the automation of an element that takes more effort is achieved, it may have greater potential impact when automated.

When it comes to the degrees of effort to understand the practice conformance elements, the result was unlike previous tables (Table 3-5 and Table 3-6). First, overall RWI is significantly lower than the level of importance and the frequency of use. This means it does not take high effort to understand the practice conformance elements. Second, the order does not have many correlations to those previous criteria (Table 3-5 and Table 3-6), whereas the previous criteria have ordinal relationships. There is a relationship between the automation of practice conformance elements and the automation of measurement of practice conformance elements. When practice conformance elements are automated, it is easier to automate the measurement of the elements as well.

Based on the results from Table 3-5, Table 3-6, and Table 3-7, an assessment of potential impact is made for automated measurement of the practice conformance elements. RWIs of each question are added to derive which element is the most amenable to be measured in automated ways (Equation (2)).

$$\text{IA} = RWI_I + RWI_U + RWI_E \quad \text{.......................................(2)}$$

- IA = Impact of automation
- RWI_I = Relative weight index for importance
- RWI_U = Relative weight index for usage
- RWI_E = Relative weight index for effort

Impact of Automation (IA) is defined as the degree to which an element may indicate conformance to a practice guideline when automated. The higher the sum of RWI is, the greater the IA is, if measurement of the practice conformance element is automated (Table 3-8). This result applies for both automation of the element and automation of the measurement of the element. That is, if the practice conformance element's IA is high, measuring that element in automated ways will also be of high value. The results were cross-checked with the Part 3 question.

Table 3-8: Potential Impact of Automation of Practice Conformance Elements

Practice Conformance Elements	RWI_I	RWI_U	RWI_E	Impact of Automation	Impact level
Schedule/Timeline (S/T)	0.95	0.88	0.66	2.5	3
Itemized List of Activities (ILA)	0.91	0.83	0.53	2.3	3
Description/Narratives (D/N)	0.88	0.80	0.58	2.3	3
Benchmark Historical Data (BHD)	0.89	0.66	0.69	2.2	2
Relevant Organization Chart/RASCI (ROC)	0.90	0.80	0.53	2.2	2
Definition/Acronym (D/A)	0.80	0.84	0.59	2.2	2
Revision Number/Date (RN/D)	0.92	0.86	0.42	2.2	2
Relevant Project Form/Checklist (PF/C)	0.89	0.74	0.56	2.2	2
Process Flowchart (PF)	0.87	0.74	0.58	2.2	2
Document Approval/Authorization (DA/A)	0.92	0.86	0.41	2.2	2
Table of Contents (ToC)	0.89	0.86	0.43	2.2	2
Relevant Chart/Graph (C/G)	0.75	0.66	0.73	2.1	1
Relevant Photo/Image (P/I)	0.84	0.75	0.37	2.0	1

The impact of automation can be between minimum 0.0 to maximum 3.0 since it is the sum of relative weight indices (RWIs). The sums were all above 2.0, indicating that all elements have somewhat high levels of impact. To analyze further, these thirteen elements were categorized into three groups (1: relatively high 2: moderate, and 3: relatively low), threshold set as 2.2 by calculating 25^{th} and 75^{th}

percentiles of data. The grouping prevents from one response changing the outcome due to the sample size. For instance when sensitivity is explored, Level 1 element cannot be Level 3 even when a participant changes a response from strongly disagree (1) to strongly agree (5) since the change affects 0.066 increase from the original value.

The grouping supports the selection of the conformance elements for the validation test of this thesis presented in Chapter 4 and Chapter 5. For the validation test, “description/narratives” and “process flowchart” elements are selected. “Description/narratives” represents the document component while “process flowchart” represents the workflow component. Details are further explained in later chapters.

3.2.2.3 Examinations of Practice Conformance Elements

To develop a concrete and specific practice conformance measurement framework, further analysis of the survey respondent’s perspectives on practice conformance elements was necessary. Analysis from the previous section strongly validated that measurements of the various elements may have different levels of impact when automated. Statistical analysis from this section further support the previous section by providing perspectives among respondent categories. Hypotheses are tested based upon the number of respondents in a category and logical judgment. Table 3-9 presents respondent categories that are used for statistical analyses.

Table 3-9: Respondent Categories

Categories	Respondent Categories (The Number of Respondents)	Respondent Categories (The Number of Respondents)
Industry Sector	Upstream & Midstream (Oil and Gas) (32)	Downstream & Chemicals (Refining, Petrochemical, etc.) (14)
Stakeholder Group	Owner (18)	EP/EPC/EPCM, Contractor (39)
Project Phase	Preconstruction (Feasibility & Conceptual Planning, Design & Engineering) (15)	Construction (36)
Role	Project Management (17)	Construction Management, Engineering Management (21)
Work Location	Site/Field-based (29)	Corporate/Home-based (32)
Age	Less than 40 (24)	40 or More (37)
Experience in Industry	20 or Less (36)	More than 20 (25)
Level in Organization	Individual Contributor (15)	Mid/Senior-level Contributor (46)

For the industry sectors, upstream & midstream (oil and gas) and downstream & chemicals (refining, petrochemical, etc.) outnumber other sectors. These two categories are compared to each other. For the stakeholder groups, owner and EP/EPC/EPCM/contractors have distinct attributes. When it comes to project phases, feasibility & conceptual planning and design & engineering are categorized as preconstruction and are compared to construction. Project management and construction/engineering management are divided based on the focus of the roles. Work location, age, experience in the industry, and level in organization are divided considering the number of respondents. The survey data is divided in such a way not only to increase the power of statistical analysis but also to make sense of the clear division. Below are the sets of respondents that are assumed to produce significant differences in results. In the following list of categories, the numbers in the brackets are the number of corresponding respondents.

To test the hypotheses, the Mann-Whitney test is applied. Among tests to find differences between two groups, such as multivariate analysis of variance (MANOVA) and logistic regression, Mann-Whitney is selected because the distributions are not normal, and the sample sizes are small. The null hypothesis (H0) is that there is no difference between the two groups. With 0.05 (95%) confidence level, when $p<0.05$, the null hypothesis is rejected. Some example hypotheses are presented below following the summary of statistic results in Table 3-10. Significant differences are described with the * mark, indicating that the criteria of the * marked categories are significantly greater than the other.

For example, there is a significant difference when comparing the perception towards the importance of the "process flowchart" between upstream & midstream and downstream & chemicals. The "process flowchart" is perceived as more important to upstream & midstream than downstream & chemicals.

Table 3-10: Significant Differences in Categorized Groups

Criteria	Category	Significant Difference * indicates significantly greater than the other
Importance	**Industry Sector**	**Upstream & Midstream vs. Downstream & Chemicals**
	Process Flowchart	Upstream & Midstream* vs. Downstream & Chemicals
	Benchmark Historical Data	
	Schedule/Timeline	Upstream & Midstream vs. Downstream & Chemicals*
	Project Phase	**Preconstruction vs. Construction**
	Process Flowchart	Preconstruction* vs. Construction
	Definition/Acronym	
	Role	**Project Management vs. Construction Management, Engineering Management**
	Document Approval/Authorization	Project Management* vs. Construction Management, Engineering Management
	Age	**Less than 40 Years Old vs. 40 Years Old or More**
	Process Flowchart	Less than 40 Years Old vs. 40 Years Old or More*
	Benchmark Historical Data	
	Participant's Level	**Individual Contributor vs. Mid/Senior-Level Contributor**
	Process Flowchart	Individual Contributor vs. Mid/Senior-Level Contributor*
	Benchmark Historical Data	
	Relevant Organization Chart/RASCI	
	Schedule/Timeline	
Usage	**Stakeholder Group**	**Owner vs. EP/EPC/EPCM, Contractor**
	Table of Contents	Owner vs. EP/EPC/EPCM, Contractor*
	Schedule/Timeline	
	Work Location	**Site/Field-based vs. Corporate/Home Office-based**
	Relevant Photo/Image	Site/Field-based* vs. Corporate/Home Office-based
	Description/Narratives	Site/Field-based vs. Corporate/Home Office-based*
	Age	**Less than 40 Years Old vs. 40 Years Old or More**
	Document Approval/Authorization	Less than 40 Years Old* vs. 40 Years Old or More
	Relevant Photo/Image	
	Experience in the Industry	**20 Years or Less vs. More than 20 Years**
	Relevant Photo/Image	20 Years or Less* vs. More than 20 Years
	Participant's Level	**Individual Contributor vs. Mid/Senior-Level Contributor**
	Relevant Organization Chart/RASCI	Individual Contributor vs. Mid/Senior-Level Contributor*
	Relevant Photo/Image	Individual Contributor* vs. Mid/Senior-Level Contributor
Effort	**Stakeholder Group**	**Owner vs. EP/EPC/EPCM, Contractor**

	Relevant Project Form/Checklist	Owner* vs. EP/EPC/EPCM, Contractor
	Itemized List of Activities	
	Role	**Project Management vs. Construction Management, Engineering Management**
	Process Flowchart	Project Management* vs. Construction Management, Engineering Management
	Definition/Acronym	
	Relevant Organization Chart/RASCI	
	Relevant photo/Image	
	Description/Narratives	
	Itemized List of Activities	
	Work Location	**Site/Field-based vs. Corporate/Home Office-based**
	Benchmark Historical Data	Site/Field-based vs. Corporate/Home Office-based*
	Relevant Project Form/Checklist	
	Description/Narratives	
	Experience in the Industry	**20 Years or Less vs. More than 20 Years**
	Relevant Chart/Graph	20 Years or Less* vs. More than 20 Years

Perspectives of participant groups are revealed. The “process flowchart” is perceived to be more important to upstream & midstream than downstream & chemicals. In the oil and gas industry, upstream and midstream refer to production and delivery whereas downstream and chemical refer to sales that deal with end-users. A “process flowchart” is used to help enforce a series of recurring tasks. The “process flowchart” may be more critical for the production and delivery procedure since the sequences of work tasks of oil and gas production and delivery are more complex than the process of the sales. The “process flowchart” element has a high potential to be automated in the form of a workflow engine. In later chapters, conformance to a process flowchart is measured in semi-automated ways.

The “benchmark historical data” is also perceived to be more important to upstream & midstream than downstream & chemicals. Generally, the upstream and midstream industry sectors have business-to-business (B2B) relationships whereas the downstream and chemical sectors have business-to-consumer (B2C) relationships since the upstream and midstream require engineering and construction. The “benchmark historical data” may be more important in a B2B relationship to encourage partnerships among corporations and to collaborate with various stakeholders. In a B2B relationship, defining and aligning the key performance indicators (KPIs) and core values are critical tasks. Once they are defined and aligned, the performance must be measured with appropriate metrics. The performance data will be stored as the benchmark historical data. Then, the practice conformance measurement can be achieved by comparing current performance measurements with historical data. Thus, there is a need for an automated benchmarking and metrics data collection program. Currently, self-reported or manually recorded evaluations exist; but automation is barely involved. In later chapters, practice conformance measurement of automated recordings of activities is addressed.

On the other hand, the "schedule/timeline" is perceived to be more important to downstream & chemicals than upstream & midstream. Conformance to a "schedule/timeline" refers to meeting deadlines and keeping track of changes. The "schedule/timeline" may be more important for downstream & chemical sector because they have a number of clients to schedule orders. To prepare products on time and satisfy clients, some downstream and chemical companies apply backward scheduling. The practice conformance measurement of the "schedule/timeline" has potential for automation, if enterprise resource planning (ERP), automated workflow engines, and integrated scheduling software systems are being used.

The "process flowchart" and the "definition/acronym" are perceived more important to preconstruction than construction. When it comes to project phases, the execution phase is easier, while the planning phase is robust. Especially for capital projects where processes are complex, various stakeholders are included, and a number of changes are expected, communication in the preconstruction phase is of importance. The reason that "process flowchart" and "definition/acronym" elements were perceived to be more important to the preconstruction phase may be related to the importance of communication in the preconstruction phase. For example, when a stakeholder is not involved in definition/alignment process in the planning phase, this may create great confusion because the stakeholder are not able to understand the language. To reduce the communication gap, definitions must be aligned and the overall process must be outlined in the preconstruction phase. The conformance to the "definition/acronym" refers to the existence of terms with correct descriptions. There is potential to automate the conformance measurement of the “definition/acronym” with toolsets, such as keyword extraction or text matching suggested by this study in later chapters.

The “document approval/authorization” is perceived to be more important to people related to project management than people related to construction management or engineering management. The “document approval/authorization” is also perceived to be used more frequently by ages less than 40 years old than by 40 years old or more. People related to project management have a different focus than people related to construction management or engineering management. For the people related to the project management, the automation of “document approval/authorization” element conformance measurement is beneficial.

The result of the “document approval/authorization” perceived to be used more frequently by the less experienced group can be interpreted in terms of responsibilities. It is mostly the less experienced group that needs to deal with and conform to the correctly approved documents. The documents may sometimes be missing or even be signed by the wrong people. Documents signed by the wrong people create confusion and delays. Therefore, automation of “document approval/authorization” has potential through automated signature verification methods along with automated document approval workflow and document version updates. Automated signature verification algorithms (Carnes, 1984; Collot et al., 1991; Mettyear, 2016) or studies (Chalechale et al., 2004; Deng et al., 1999; Gideon et al., 2018; Rudyi et al., 2019) can be used for this purpose. While some studies are based on mathematic equations, others use image processing technology (Hussein et al., 2016) for signature verification. Optical character recognition (OCR) technologies, which are also related to the automated conversion of handwritings into machine-readable texts, are discussed further in later chapters.

The “relevant organization chart/RASCI” is perceived as both important and frequently used by the mid/senior-level contributor group. There are three implications of this. First, the results prove that the survey participants responded with integrity since the level of importance and the frequency of use are related. Second, this information has an additional value due to higher expertise of respondent categories. It is likely that higher level contributors have insights and bigger visions whereas the lower level contributors focus on details that one is involved in. Therefore, it implies that the impact of automation may be higher than other elements. Third, this information allows an auditor to separate the employees into groups and ask unique questions during interviews. For example, the auditors may ask the holistic process- or organization-related questions, such as the “process flowchart,” “benchmark historical data,” “relevant organization chart/RASCI,” and “schedule/timeline” for the more experienced group while asking task-related specific and detailed questions for the less experienced group. This will save time for auditors, and provide them with more specific and accurate results.

Concerning “relevant organization chart/RASCI,” software packages such as Microsoft Visio provide a service to create an organization chart automatically from employee data. By simply entering employee names, unique identifiers, and to whom the employees report, the organization chart is created. The automation of the “relevant organization chart/RASCI” implies the potential of the automation of the conformance measurement of the element. There is a potential for the automatic detection of the appropriate roles and responsibilities that are missing or mismatched in the “relevant organization chart/RASCI.”

The “table of contents” and “schedule/timeline” are perceived to be used more frequently by EP/EPC/EPCM/contractor than owner. One of the tasks for EP/EPC/EPCM/contractors is to define their scope clearly to complete the construction in time. If projects are delayed, they incur the penalty more directly. Therefore, it is reasonable for the EP/EPC/EPCM/contractors to use the “table of contents” and the “schedule/timeline” more often and take time to build suitable ones. The conformance measurement of “table of contents” have a potential for automation with text matching toolsets suggested by this study in later chapters.

The “relevant photo/image” is perceived to be used more frequently by site/field-based employees than corporate/home office-based employees. The “description/narratives” are perceived to be used more frequently by corporate/home office-based employees than site/field-based employees. When it comes to frequencies of use, distinctive perceptions between the site/field-based and corporate/home-office based employees are apparent. Due to their characteristics and functions, whereas the site/field-based employees have a strong sense of presence, intuitiveness, and familiarity with “relevant photo/image,” documentation (“description/narratives”) is one of the main tasks for the corporate/home office-based employees. For employees from both work locations to reduce communication gaps and conflicts, they need to acknowledge how the other works and search for ways to efficiently collaborate. If they possess technologies such as the advanced reverse image search technologies that translate the elements into their own strengths in automated ways, it may be a great asset that saves time and confusion.

The image matching or computer vision tools have the potential to match other “relevant photo/image” as a part of automated practice conformance measurement. For the “description/narratives,” text matching toolsets suggested by this study in later chapters can be used to automatically detect matching between two related bodies of texts. Furthermore, for the “relevant photo/image,” image reverse search toolsets may translate images into relevant words. This automated translations between “relevant photo/image” and “description/narratives” would benefit both employees working in separate locations.

When it comes to a checklist, two levels namely the management-level (i.e., owners) and the practice-level (i.e., EP/EPC/EPCM/contractors) are associated. These two levels have unique points of view. Whereas practitioners (EP/EPC/EPCM/contractors) need checklists to work in an organized order, managers (owners) need checklists so they do not miss any important issues. The scope of interest and focus are dissimilar. Moreover, because of the specializations, typically, EP/EPC/EPCM/contractors have similar roles and tasks for various projects. Therefore, the “relevant project form/checklist” may take less effort to understand. On the other hand, every project differs for owners; thus, creating the "relevant project form/checklist" and understanding it may take more time and effort. Likewise, when it comes to the "itemized list of activities," assigning the specific work that is similar to previous projects is less difficult than listing the items for projects that have challenging dynamics.

The "relevant project form/checklist" has the potential for automation of practice conformance measurement. With current technology, when the "relevant project form/checklist" is created electronically, the creators typically have the ability to force the participants to respond to prevent them from skipping the questions. This feature reduces human errors and enhances practice conformance with semi-automation. For survey questionnaires that require manual recordings, this function may be useful. When the recording can be completed by machine after either human or machine completing the task with the integration of the internet of things (IoT) or sensors, practice conformance measurement can also be semi-automated.

The age, the experience year, and the level of contribution are related. The "relevant photo/image" is perceived to be used more frequently by ages less than 40 years old than by 40 years old or more; and more frequently by 20 or less years of experience in the industry than by more than 20 years; and individual contributors more than mid/senior-level contributors. When it comes to the experience level, lower level contributors may not be able to understand "description/narratives" due to their lack of experience. They require at least some visualizations to imagine the "description/narratives." For higher level contributors, on the other hand, written documents are enough for understanding.

The "relevant chart/graph" is perceived to take more effort to understand by 20 or less years of experience in the industry than more than 20 years. The less experienced group may have not completed as many tasks related to the "relevant chart/graph." However, the chart and graph are graphical representations of data. Thus, when it is partitioned in automated ways, it will aid understanding by the less experienced group. When the visualization ("relevant chart/graph") is created automatically in real-time, it is easier to track the current data and compare against the past data. Conformance is achieved when the real-time data matches the benchmark. The data collection system (DCS) can integrate with the internet of things (IoT) or sensors to facilitate the data collection process. By synchronizing the real-time data with "relevant chart/graph," automation is applied and conformance measurement is achieved.

Overall, the results and analyses demonstrate that positions and stances impact one's perspectives on capital project practice conformance elements. Some conformance elements are perceived as more important, more used, and take more effort to understand than others and each conformance element has the potential to be measured in automated ways. Findings in this section assist companies, consultants, and auditors to prioritize the automation and automated conformance measurement of practice conformance elements. The findings also guide software developers and engineers to automate parts of conformance elements or the conformance measurement based on the information of the potential automation for each conformance element.

3.3 Discussion and Conclusions of the Practice Conformance Model

A practice conformance model is a representation that describes an operational adherence to practice guidelines (Figure 3-2). By dividing into four components, namely documents, workflows, people, and actions, the practice conformance model can be used to measure practice conformance. The practice conformance components can be measured by checking their existences and their correctness. Moreover, the practice conformance components are decomposed into practice conformance elements to practically measure the degree of conformance and potentially automate the practice conformance elements and the measurement. The practice conformance elements are validated for industrial applicability. Through the rounds of meetings and a survey, the practice conformance elements are validated.

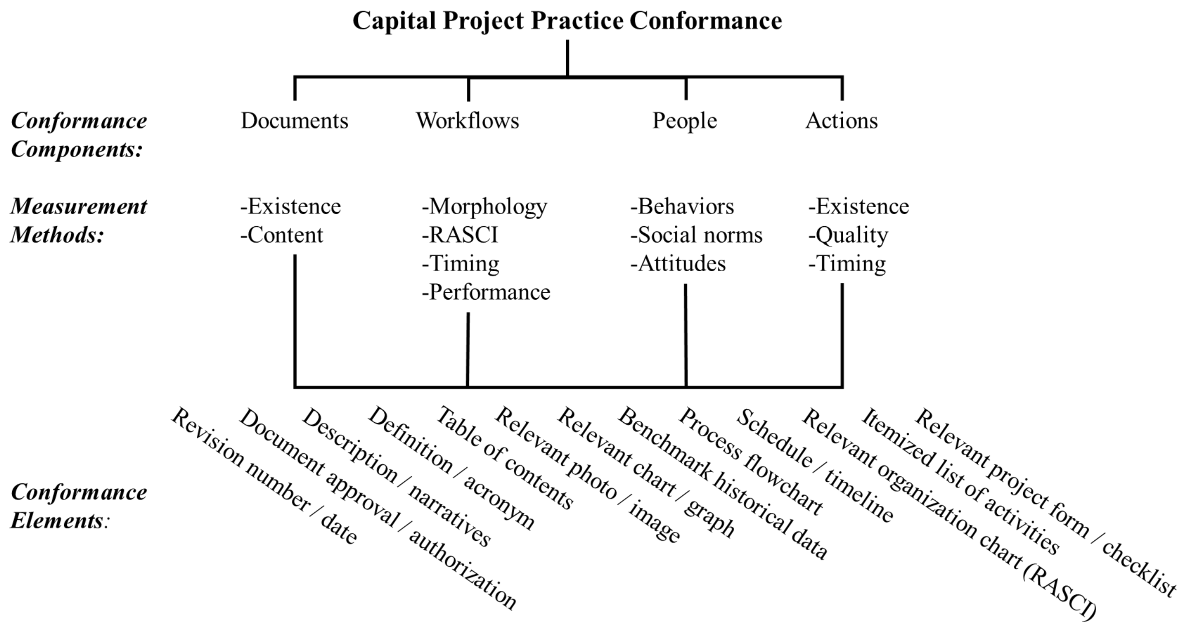


Figure 3-2: The Capital Project Practice Conformance Model (Final)

Furthermore, the practice conformance components and practice conformance elements are not independent but interrelated. For instance, every practice conformance element (i.e., all thirteen practice conformance elements) can be an element of the document component if it is in an appropriate form. Some elements, such as "benchmark historical data," "process flowchart," or "schedule/timeline" are interconnected with the workflow component as their morphology, timing, and performance can be measured for the assessment of these elements. "Relevant organization chart (RASCI)" is related to the people component while "relevant project form/checklist" is correlated with action component since forms and checklists are records of actions.

Specifically, the practice conformance elements may be used when distinguishing the types of information. For example, in Chapter 4, there is an "information type detection" module in the document conformance measurement framework. In this module, the existence of the practice conformance elements may be checked. When important elements such as schedule and timeline are missing, the document is recorded as non-conformance.

The impact of practice conformance elements automation can also be foreseen from this study. The greater the importance is, the more it is used, and the greater the challenge there is to understand the practice conformance elements, the automation will have a greater impact on the industry. The impact of automation may be useful for not only auditors but also software developers who have capital projects clients.

With the result statistics in Section 3.2.2, customized methods are possible to establish. The practice conformance elements are what constitute the framework. This means that the framework, methods, and toolsets for different types of capital projects can be tailored differently. Some companies may be willing to invest more in automation than others. Some may prioritize accuracy over time. The practice conformance elements and the analysis will aid in tailoring the customized methods. A few method examples are introduced and tested in Chapter 6. Specifically, the statistics regarding the level of importance, the frequency of use, and the degree of effort to understand the practice conformance elements from Section 3.2.2.3 provide information on the approaches according to the characteristics of a company. Table 3-11 summarizes the potential for practice conformance element automation and the advantages for practice conformance element measurement when automation is achieved.

Table 3-11: A Summary of Potential Practice Conformance Element Automation and Advantages

No.	Practice Conformance Element	Potential for Practice Conformance Element Automation	Advantages in Practice Conformance Measurement
1	Schedule/Timeline	Enterprise Resource Planning (ERP), automated workflow engines, and integrated scheduling software systems keep track of activities and resources to meet the deadlines	Easier to measure practice conformance since activities are recorded in automated ways; doesn't require external data or large data
2	Process Flowchart	Automated workflow engines enforce the process flowchart and alert when there are violations; process mining software analyzes the pattern of the real data	Easier to measure practice conformance since the process flowchart is in machine-readable form (e.g., BPMN) and data are recorded in automated ways; comparison between the process flowchart and real data is achieved automatically
3	Relevant Organization Chart (RASCI)	An interface management system (IMS) creates an organization chart that involves automatic reporting structures; gives authorities to create, view, edit the documents or any deliverables	Having a well-defined IMS implies practice conformance
4	Itemized List of Activities	Automated keyword extraction toolsets simplifies the list of activities into a list of keywords visually and automatically	Having the list of keywords that might have been overlooked implies practice conformance
5	Description/Narratives	Automated text matching toolsets detect similarities and differences of the description/narratives	Not having necessary clauses implies non-conformance
6	Document Approval/ Authorization	An automated signature verification method (image processing, optical character recognition (OCR)), automated document approval workflow, and automated document version updates allows the employees to work on the latest approved versions of documents	Having a well-defined system that identifies the signatures, updates correct versions after approval to relevant employees implies practice conformance
7	Table of Contents	Automated keyword extraction toolsets remove unnecessary words and text matching toolsets detect matching between remaining words	Not having necessary matching keywords implies non-conformance
8	Relevant Project Form/ Checklist	The integration of the internet of things (IoT) or sensors enables automated recording of data in project form/checklist; image processing, optical character recognition (OCR) and/or forms/checklist created in an electronic format check whether all questions are answered and notifies when questions are not answered	Automatic recording reduces human errors and subjectivities; notifying unanswered questions is a form of conformance

9	Benchmark Historical Data	An automated benchmarking and metrics data collection program collects benchmark historical data with electronic survey questionnaires and voice recognition (voice-to-text)	Recording the data electronically reduces human errors; identifying the best practices from the benchmark historical data is a form of conformance
10	Revision Number/Date	An automated versioning toolset allows to find the latest version of the documents by finding the latest timestamps (save); image processing when the version number or date is located in the same spot of documents	Automated correct versions update implies a form of conformance
11	Definition/Acronym	Automated text matching toolsets detect matching between words	Non-existence of necessary words implies non-conformance
12	Relevant Chart/Graph	The data collection system (DCS) integrates with the internet of things (IoT) or sensors; real-time charts or graphs are created with automatic analysis on changes	Data are recorded in automated ways; human errors and subjectivities are reduced; automated DCS implies higher chance of conformance
13	Relevant Photo/Image	Reverse image search technologies; image matching and/or computer vision toolsets	Automation implies better communication between site/field-based and corporate/home office-based employees, implying some degree of conformance

For further chapters, the document and workflow components are emphasized since these two components are visible, tangible, and measurable. They may also represent the people and action components because the people and action components are who create and follow the document and workflow components. From the literature review, the document and workflow components were categorized as two fields, “automated compliance checking” and “process mining.” However, this study differs from previous studies when it comes to approaches.

The document conformance measurement approach is in relation to automated compliance checking. There have been some studies on automated semantic compliance measurement (Yurchyshyna & Zarli, 2009; Salama & El-Gohary, 2011; Salama & El-Gohary, 2016; Zhang & El-Gohary, 2016). These studies examined the characteristics of words in the body of texts and their relations. For example, when negations are added, they deliver an opposite meaning in context. Whereas these studies can validate the need for semantic compliance checking, because there are too many exceptions in the real-world texts, the tests are limited to proof-of-principle demonstrations. Though they support the importance of automated conformance measurement, such past studies lack industrial applicability. On the other hand, this study approaches automation with a model, framework, and toolsets. This approach is practical and flexible with the conditions of companies when it comes to adoption. Moreover, the approach is holistic since not only the texts, but the entire documents are the scope of practice conformance measurement. Commercial tools suggested in this study are available for any company.

The workflow conformance measurement approach is in relation to process mining. The process mining field is a relatively new field than the automated compliance checking field. Hence, early-stage technology lowers accuracy and usefulness. In the construction industry, few scholars or practitioners know about the field. Due to the current pioneer stage of theories, principles, and technologies, this study had difficulties in comparing methodologies to other studies. However, the importance of the workflow component is universal, yet the framework to measure it is lacking in the construction domain. Considering that recordings are completed in automated ways, process mining reduces human errors. Although the process mining approaches are perceived as useful conceptually, not all concepts and methods provided from this study are adopted from the process mining field.

To establish the practice conformance model presented by this study, opinions from the industry and academia have been collected. The model is balanced since it includes both perspectives of the industry and academia. That is, academia will grow by developing the frameworks and methods to automate the practice conformance components in semi-automated ways. The industry will actively use the suggested framework and available toolsets to increase productivity.

The semi-automation of practice conformance component measurement through the practice conformance elements will not only facilitate the audit process but maximize the efficiency of the construction industry. The results of this analysis will assist an auditor who needs to decide the scope and methods of an audit. The auditor will benefit when the selected elements can be semi-automated.

The practice conformance model is the foundation for the framework and toolsets that are further explained in later chapters.

Chapter 4 Semi-Automated Capital Project Practice Conformance Measurement Framework

In the previous chapter, the document and workflow conformance components were demonstrated to be tangible, and their conformance measurable. The other components, people and actions are also as important, and must not be ignored. However, people and actions are involved in documents and workflows since they are the participants who create and follow these components. Because all components are interrelated, and they are not independent of each other, by measuring two components, practice conformance measurement of the other components are also partly accomplished. In this chapter, methods to measure document and workflow conformance are presented. Step-by-step processes and flowcharts are intended to help understanding and to encourage the adoption of the framework. Figure 4-1 is a functional decomposition diagram that explains the capital project practice conformance measurement framework. It includes functions that are needed for practice conformance measurement.

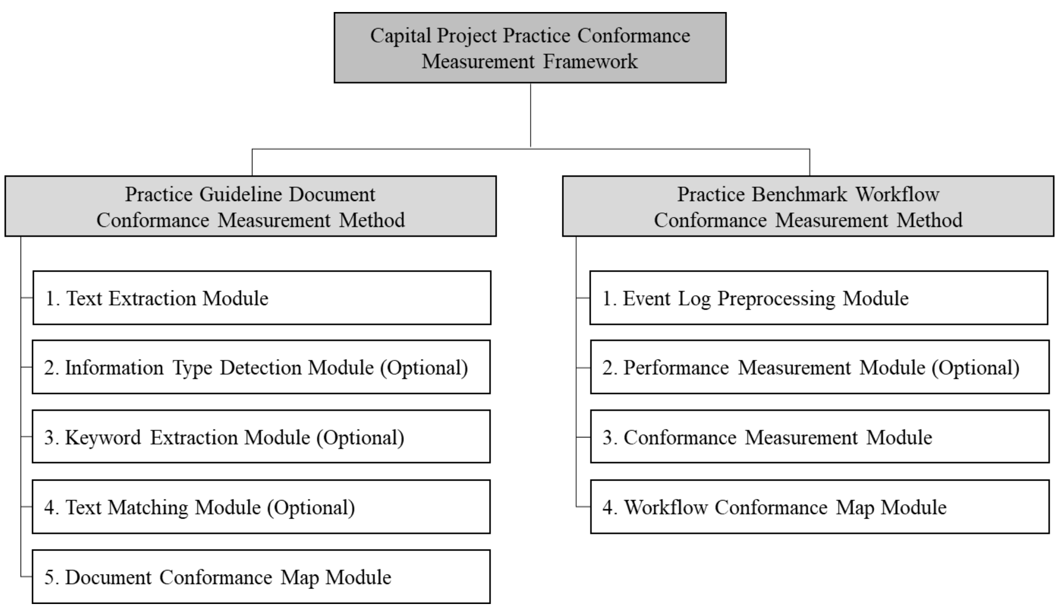


Figure 4-1: Functional Decomposition Diagram of Capital Project Practice Conformance Measurement Framework

The goal of these processes is to reduce cost and time yet to increase accuracy. There are optional modules that can be skipped. By including or excluding these optional modules, auditors will find the best process to achieve this goal. The orders are not fixed either. Auditors can decide which sequence works the best and apply it to their own situations. These processes will not only increase the productivity of the auditors but ultimately encourage companies to conform to practice guidelines and benchmark workflows.

The capital project practice conformance measurement framework suggested by this study is simply divided into two processes (Figure 4-1). In these processes, some automation is involved. First, to measure document conformance against practice guidelines, five modules may take place. When a document conformance map is finally created, it will assist auditors to determine whether the corporate documents are conforming to the practice guidelines. The information type detection, keyword extraction, text matching modules are optional and depending on the definitions of practice conformance they may be used accordingly. However, adding modules typically increases total time and costs. Thus, the auditors must consider the trade-off between the level of accuracy and time and costs.

Second, to measure workflow conformance against benchmark workflows, workflow conformance measurement must be completed producing a workflow conformance map as the result. When the performance is measured, the benchmark workflow can be improved according to the output regarding conformance and performance. When necessary, the event log preprocessing module may be taken. Event log preprocessing is necessary when the software packages do not run with the raw data.

Note that this overall framework is not how the actual manual audit process works. Every auditor or audit team has unique strategies in a manual audit. They have accumulated their knowledge and experiences in sampling and detecting practice conformance. Therefore, they may not follow this framework at all. Whereas it is difficult to standardize how manual audit works, this structured framework is intended to gain meaningful practice conformance results in semi-automated ways.

4.1 Capital Project Practice Guideline Document Conformance Method

A methodology for measuring practice guideline document textual content conformance was developed and is described in this section. Its use is demonstrated as well, and its potential efficacy is assessed in Chapter 5 and Chapter 6. Given the preceding development of functional decomposition of a capital project practice conformance measurement framework, functional requirements must follow. Such requirements are identified for the development of a practice guideline document conformance measurement. Figure 4-2 describes the specific functional requirements to achieve document conformance measurement. There are five modules, and within a module, there are some requirements that can be accomplished either manually or automatically.

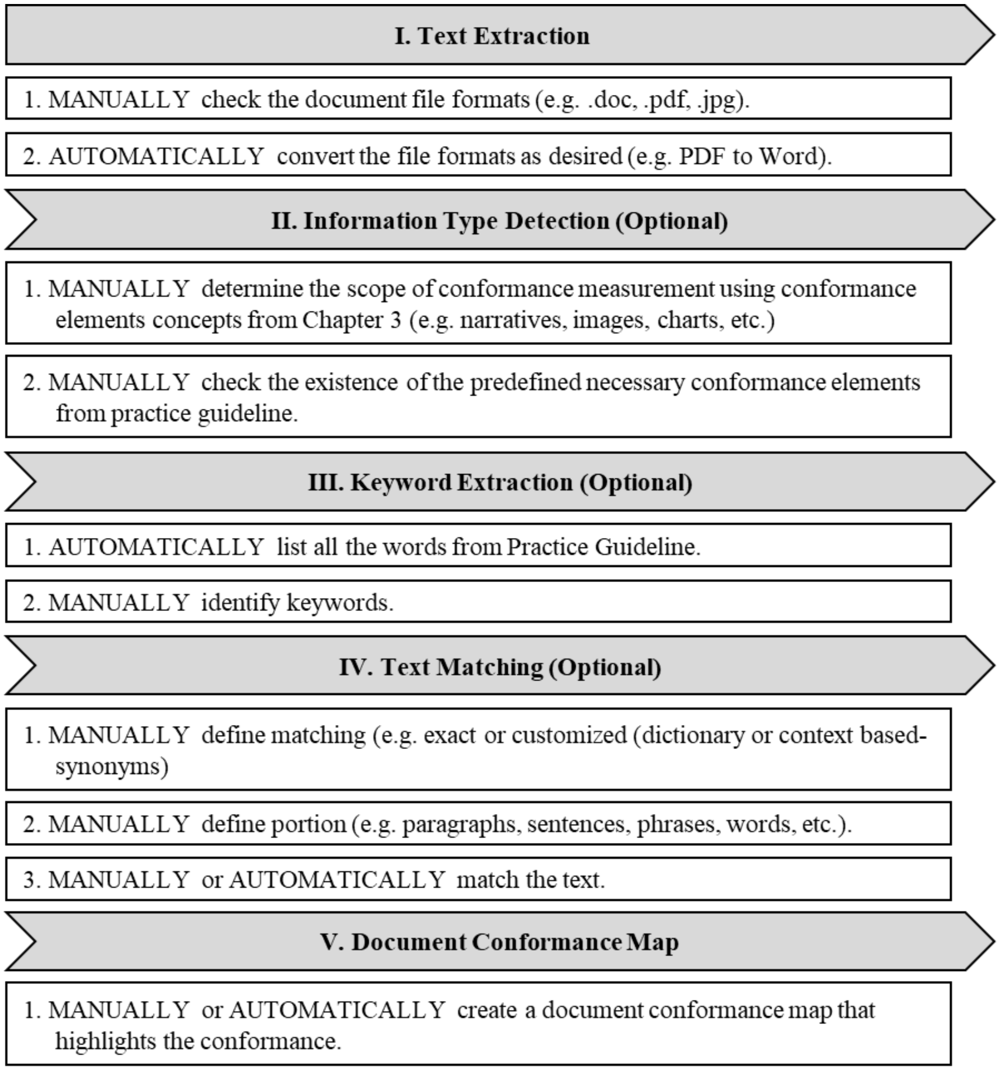


Figure 4-2: Functional Requirements of Document Conformance Measurement

Documents can be in many formats from electronic file formats (PDF, DOCX, XLSX, PPTX, etc.) to web content or even in physical paper formats. Forms of documents are not a critical issue when a manual audit takes place. However, to automate the process, the machine needs to be able to read the documents. This first module, text extraction, is the preprocessing module that is required unless the documents are in electronic form and the file formats are processable by the machine. It is generally the best to keep original format since there are risks associated with converting the file formats such as information loss or errors.

The second module, information type detection, is intended to allow broader-level conformance measurement and one-to-one comparisons. Broader-level conformance refers to finding the existence of necessary conformance elements. By utilizing the practice conformance elements that are derived from Chapter 3, this comparison is possible. Moreover, for one-to-one comparisons, this module can be followed by other optional modules, such as keyword extraction and text matching. By setting the information types (e.g., practice conformance elements) apart, matching information types are expected to produce the conformance measurement results with higher accuracy. This module (Module 2) is optional and if desired, it can exist after Module 1, Module 3, or Module 4. If the document has only one type of information type, this step is unnecessary. This step is mainly for complex documents with multiple practice conformance elements. Moreover, even if the document has more than one information type, the auditor can simply ignore the information type in the beginning. This is when Module 2 (information type detection) goes after Module 4 (text matching). That may take less time. However, there are instances where including this module (Module 2) in the beginning, increase accuracy. It takes extra time, but higher accuracy is achieved. For example, if there is a chart regarding the drug concentration level in practice guidelines, the first step is to detect whether it exists in corporate documentations. The second step is text matching to find its correctness. The order can be reversed, but, when there are many other elements, the order may affect the results.

The third module, keyword extraction, is also optional and can be included according to the definition of practice conformance. It also depends on time and budget. When sampling is necessary due to the time constraints, when the budget is low, this module will assist auditors to find and check important portions of texts. For example, if the word “tamper” is considered a keyword, the existence of the keyword can be checked. Next, to be more accurate, auditors may check the context where and how this word is used. If the context is allowing “tampering,” when the practice guideline is not, the corporate document is not conforming with the practice guideline. This module (Module 3) requires professional auditors. Listing the words can be completed by software packages, but professional knowledge and experiences is required for identifying the keywords. Thus, it may increase costs. Keywords can be indicators of where the relevant content is. This module must be followed by module four, text matching, since without matching, identifying the keywords from practice guidelines does not serve any purpose. However, this module can also be executed after the text matching module (Module 4). After text matching, there may be non-matching portions. When keywords exist even though the portion does not qualify as matching, this piece of information indicates that there are higher chances of conformance. This is when the text matching module occurs twice. Again, this module may increase time, but the accuracy of the results may be worth the time.

The fourth module, text matching, is one of the most critical modules for the document conformance measurement. Matching does not guarantee conformance. However, because matching is defined as an identical portion of texts, it helps where and what to investigate (Section 6.1.1). By executing this module (Module 4), the auditors can find out whether the necessary clauses from the practice guidelines are in the corporate documents. The auditors can notice differences in numeric values. This may significantly reduce the time for processing the documents.

The last module, the document conformance map, is a result of previous modules. A conformance map is an illustration of conformances and non-conformances that guides auditors by highlighting the portions that are in conformance. It is a visualized guideline that may assist auditors. An auditor can use the conformance map to decide which parts of documents to focus on. For example, when a mandatory clause of contract or an identical disclosure statement that a company must include is missing, the conformance map indicates its absence. The auditor may begin an investigation with this piece of information. Such a map is developed and illustrated in the following chapters.

These modules from Figure 4-2 have both manual and automatic functional requirements. Some modules are optional while others are essential to achieve this process. To reduce the time, one suggestion is to follow Module 1, Module 4, Module 2, and Module 5. Module 1 can be skipped if the document formats are compatible with the software. Manual parts that are completed in Module 3 are to enhance accuracy but may take additional time. In summary, the types and sequences of the modules must be selected with considerations. Skipping modules is a way to reduce costs and time, but, to increase the accuracy, at least one module from module two to four must be selected.

4.1.1 Text Extraction Module

Text extraction is the task of extracting text from documents so that they are machine-readable. The texts from practice guidelines and corporate documents need to be extracted for later modules. The term text extraction comes from information extraction from the text mining domain. Information extraction is the task of automatically extracting structured information from unstructured/semi-structured machine-readable documents. Likewise, the text extraction module aims to distinguish texts from images for machines to recognize the text letters. There are documents created in file formats, such as .txt, .docx, .pdf, .pptx, and .jpg formats. These original electronic file formats that have been used may not be compatible with keyword extraction or text matching software packages for semi-automation. Typically, commonly used file formats such as PDF or Microsoft Office are compatible. However, an image file format such as .jpg cannot be processed for keyword extraction or text extraction module. Thus, in the text extraction module, file converter software packages are used for automation. When a document is in an incompatible format or if the text from the document is not recognized by the machine automatically, the text extraction module is necessary. However, if not, it is not a recommended module due to the information loss or translation errors. Such translation errors include hyphens or spaces and they are explained in Chapter 5 in more detail.

Another issue that can be confronted in the construction domain is handwritten documents or any documents that involve non-electronic texts. The manual audit takes no extra effort in this case.

However, for computers to process any non-electronic texts, they need to be translated into machine-encoded text with optical character recognition (OCR). Optical character recognition is a technique of digitizing printed texts so that they can be electronically edited, searched, and stored. This step is also crucial for text extraction when necessary. Though the concept of OCR software can simply be explained as converting handwritten documents into machine-readable texts, the principles underlying are complex and require machine learning. The similarities of character shapes, the overlaps, and the interconnection of neighboring characters of cursive handwriting complicate the problem (Arica et al., 2001). Handwriting recognition requires training and recognition phases. While holistic strategies consider the whole word as a target, analytic strategies require segmentation of the word by stroke or character level. In the training phase, global parameters such as stroke width/height, slant angle, lower-upper baseline are estimated with training images. Then segmentation takes place followed by feature extraction. In the recognition phase, a test image may undergo a similar process as the training phase. The recognition phase is affected by the training phase when features are extracted, and the words are recognized. The recognition phase utilizes the training phase data. Not all OCR tools have the same level of accuracy. Even within a tool, the accuracy varies due to the resolution or direction of an image.

4.1.2 Information Type Detection Module

The information type is a category of elements from documents and they can overlap with practice conformance elements. That is, practice conformance elements may be the types of information. Information type detection refers to distinguishing the categories for comparisons of the corresponding categories. For example, a "process flowchart" element is a different information type than the "schedule/timeline" element. Hence, the "process flowchart" is recommended to be compared to another "process flowchart" for an accurate comparison. This module is a manual module that is added for higher accuracy in practice conformance measurement, and may also be completed after text extraction, keyword extraction, or text matching module. It requires human judgment whether to include this module or not due to the information loss risks. For instance, if there are two documents with images in different locations, when the images are removed, the locations of texts are mixed. This process of removing the images creates confusion for computers to process the texts without proper alignment.

To match corresponding information types, capital project practice conformance elements derived from Section 3.2 can be used. If a necessary practice conformance element from a corporate document is missing from a practice guideline, it can also be recorded as non-conformance, and reported to the document conformance map module. For example, if the "schedule/timeline" element is missing, when timing and meeting the deadline are the essential factors, then, auditors can notify the company that the essential element is missing in the project management plan document. This correlates to Section 3.2.2, because what each company should focus on is statistically analyzed based on their characteristics. As an example, if the auditor is working for upstream and midstream corporations, based on the statistics, it is wise for them to have "process flowcharts" and "benchmark historical data," whereas for downstream and chemical corporations, "schedule/timeline" is essential. This module ensures what elements must exist in the documents to increase accuracy in practice conformance measurement. Another example is preconstruction and construction phase documents. According to Table 3-10, process flowcharts and definitions are typically more important for preconstruction than construction phase documents. Thus, it is advised for the auditor to search for these elements in preconstruction documents.

When an auditor desires to investigate further than merely checking the existence of information types or when the auditor concludes that information types are interrelated, it may be useful to distinguish between texts and images. For instance, an "itemized list of activities" element may be included in the "description/narratives" element. "Schedule/timeline," "process flowchart," and "relevant organization chart/RASCI" elements may be combined. In such cases, due to the scattered information, a simple distinction between texts and images may be practical. In Table 4-1, information types are categorized as either image, text, or both. The X mark represents that the element includes relevant images or texts.

Table 4-1: Inclusiveness of Texts and/or Images in Practice Conformance Elements

No.	Capital Project Practice Conformance Element	Include Texts	Include Images
1	Schedule / Timeline	X	X
2	Process Flowchart	X	X
3	Relevant Organization Chart / RASCI	X	X
4	Itemized List of Activities	X	-
5	Description / Narratives	X	-
6	Document Approval / Authorization	X	X
7	Table of Contents	X	-
8	Relevant Project Form / Checklist	X	X
9	Benchmark Historical Data	X	X
10	Revision Number / Date	X	-
11	Definition / Acronym	X	-
12	Relevant Chart / Graph	X	X
13	Relevant Photo / Image	-	X

The distinction between texts and images can be achieved by keyword extraction or text matching software automatically. Therefore, it may be easier to implement this module at the end of the framework just before creating the document conformance map (Section 4.1.5). However, if the auditor intends to compare a specific element, the information type detection module will save time and costs by comparing selective elements. For example, the "table of contents" of a project management plan may be the only element that needs to be compared. By either removing all other elements or simply extracting the "table of contents" element, not only the scope is reduced but also accuracy increases due to fewer distractions and confusion for computers.

Once information types are detected and matched, they are ready to move on to the next, keyword extraction or text matching module. If the auditor finds that this module is enough for overall document conformance measurement, the auditor may move on to module five, document conformance map. If the auditor chooses to compare elements that only include an image, image comparison methods or image reverse search can be an option (Wilson et al., 1997). However, image comparison methods are not the scope of this thesis.

Although there is no current technology that can be applied to support the information type detection module automatically, to separate text from images, the software packages used for the text extraction can be utilized.

4.1.3 Keyword Extraction Module

Keyword extraction is the task of extracting the most relevant words from texts. Keyword extraction is used in text mining, information retrieval, and natural language processing (NLP) for index generation, query refinement, and text summarization. There are many ways to emphasize a word or concept in a document. Frequently appearing words may be keywords. The words may also be in a bigger font, italicized, bold, or underlined. The strength of the keyword extraction module in this framework is that the keywords can be indices to point to the important portions of texts. With this module, some important portions that may have been missed due to the lack of time can be acknowledged. This module also may increase accuracy for measuring practice conformance. It is also a great module for sampling out important portions when there are time restrictions to check whole documents. This module can be after the text matching module to be even more specific and accurate. Non-matching portions that contain some keywords have higher chances of being conformant. This module requires relatively more professional judgment than other modules.

Some studies can be used for keyword extraction such as simple statistics approaches to more advanced linguistic approaches or machine learning approaches. The benefits of simple statistics approaches are that data training is not required, and methods are language and domain independent. Linguistic approaches use lexical, syntactic, semantic, and discourse analyses. Machine learning approaches are divided into supervised and unsupervised approaches. They require training such as manual annotation in the training dataset. Both linguistic and machine learning approaches are complex and machine learning approaches are demanding and time-consuming (Beliga, 2014). Term Frequency-Inverse Document Frequency (TF-IDF) is an example of simple statistics approach. In this approach, importance of a keyword increases in proportion to the frequency of occurrence of a word in the document but is countered by the total number of documents having that word (Siddiqi and Sharan, 2015).

4.1.4 Text Matching Module

The text matching module is the module to achieve the main objective of practice conformance measurement. It is a module that can be completed manually or automatically. Depending on the budget, time, and desired accuracy, the auditor can select ways to complete this module.

When it comes to the text matching module, defining the matching is one of the most critical steps. The definition of matching can vary, and potential definitions are present with case studies in Chapter 6. Depending on the definitions made by an auditor, the results may vary. Matching may sometimes be defined as an exact match of the body of texts. That is, slight deviations will be considered as non-matching. Other times, paraphrased texts or dictionary- or context-based synonyms still may be considered as matching. Several software packages with proprietary algorithms can be used for this purpose and are described later.

However, the output that is returned by the software packages is limited. Not all definitions can be considered due to technical limitations. Especially, to recognize the context-based synonyms, machine learning is essential. This is future work for text-matching software companies to improve their software for domain-specific documents.

When it comes to portions that the auditors are comparing to, it is advised not to be too complex nor to be too simple. For example, paragraph-level is convenient, because if two paragraphs from two documents match exactly, it is highly likely that the corporate documents followed or at least considered the practice guidelines. Although it is early to conclude that the entire corporate document is conforming to the practice guidelines, the paragraph-level provides convincible evidence to make such a claim. Sentence-level or word-level can also be used; however, it is harder to claim that the corporate documents referenced practice guidelines because a couple of matching results could just be a coincidence.

Text matching becomes more accurate with natural language processing (NLP). Natural language processing is a field where human languages are read and interpreted by computers. With NLP, large data are trained by computers to understand the semantic meaning of texts in contexts using inference. Even when two portions are written in completely different words, after the training, the computers will recognize them as same meaning and identify them as conformant portions. Generally, NLP requires machine learning, which can be divided into supervised and unsupervised learning. Typically, when supervised learning is selected, labels are assigned to words. Then, manually designed features are extracted. Next, the words are classified into previously designed features empirically with trials and errors.

Some past studies use a multilayer neural network to reduce manual preprocesses (Kaur et al., 2014). This is related to unsupervised learning. Those previous studies attempt to develop the NLP system that begins with the word level, moves on to sentence level, and then to the context and the overall environment and domain. Word level is to determine the morphological structure and nature such as part-of-speech and meaning of the word. The sentence level is to determine the word order, grammar, and meaning of the entire sentence. Finally, context level is to relate to other words or sentences in the given context (Bengio et al., 2007).

4.1.5 Document Conformance Map Module

A document conformance map visually illustrates where the conformance exists. This module can be the by-product of the previous modules, such as information type detection or text matching. The results from the information type detection module regarding the non-existence of essential information types (e.g., practice conformance elements) are found manually. The results from the text matching module are completed manually or in automated ways with software packages. Highlighting the portions of texts that match practice guidelines not only indicates the existence of critical portions but also pinpoints differences in numeric values and subtle changes. Figure 4-3 illustrates the whole process of producing the document conformance map. From the text extraction module to the document conformance map module, all modules are included. Based on the auditor’s judgment, modules can mix and match. This flowchart is designed for the semi-automated methods and it has gaps from how auditors manually work in reality. Gray boxes are ones that can be completed semi-automatically.

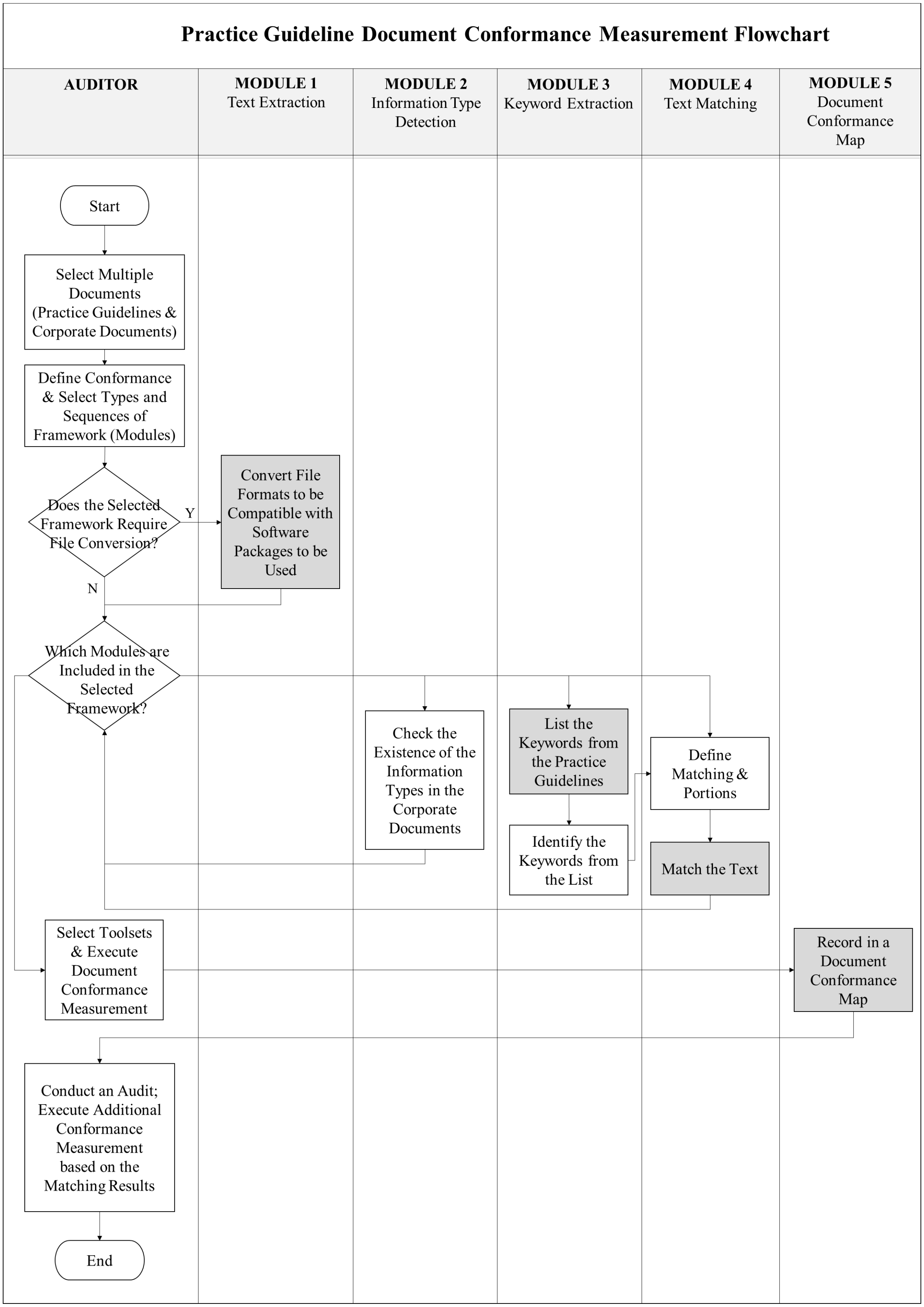


Figure 4-3: Swimlane Flowchart for Document Conformance Measurement

Conformance and matching are distinguished here since the non-matching portion can be in conformance. As discussed in Chapter 3, document conformance depends on its existence and content correctness. Matching simply finds out what is identical and what is not. When identical, it can be defined that it reached conformance (Section 6.1.1). However, if there is a slight difference, that subtlety can be a source of non-conformance. Conformance may not reached if the subtle differences are critical issues. Thus, conformance refers to the adequacy of the corporate documents whereas matching refers to finding what is identical or close to identical. Conformance is the end judgment while matching is considered as a process.

For example, consider the following practice guideline by the consumer product safety commission: "all residential garage door operators must have an inherent reversing mechanism capable of reversing the motion of a moving garage door within 2 seconds, to reduce the risk of entrapment." Some companies may have the same clause as above, while some may change words such as "motion" into "signal" to avoid the exact match or "2" to "1.8" seconds to overachieve the goal. Although the company is conforming to the practice guideline, it may not necessarily be matching when matching is defined as an identical clause.

To create a document conformance map, two documents are selected. Then, practice conformance must be defined and according to the definition of practice conformance, methods must be selected. When the methods are selected, note that their sequences make results vary. Next, there are a few questions that need to be asked to tailor the process. The text extraction module (Module 1) question would be whether the documents need file conversion. The information type detection module (Module 2) and the keyword extraction module (Module 3) questions are simply to check whether they are included in the methods. If they are not included, the auditor can move on to the next module. Notice that there is a feedback loop that allows different combinations of modules. For each module from the information type detection module (Module 2) to the text matching module (Module 4), some guideline questions can be applied. For example, if the answer is yes to the information type detection module (Module 2) question "Is the corporation document missing any essential information type?", this non-existence can be considered non-conformance. If the response is no, existence can be evidence of conformance but may require additional conformance measurement methods. In this case, the auditor can follow the next modules (keyword extraction module (Module 3) and text matching module (Module 4)). Another example question addressed in the flowchart is when utilizing the keyword extraction module. When this module is included, the definition of practice conformance highly likely is related to the keywords. This module must be followed by text matching module (Module 4). The text matching module (Module 4) is not required depending on the definitions of practice conformance. For example, the information type detection module (Module 2) may be enough considering the resources and deadlines. The text matching module (Module 4) can exist independently or after other modules. The text matching module (Module 4) is the module to check to match and finalize the results for the last module, document conformance map (Module 5). Depending on the definitions and methods that are selected, the conformance map is created in the last module. It may simply return conformance/non-conformance or add percentage values such as how much the corporate document conforms to the practice guidelines. The meaningful data would be the map that shows why it is conforming or non-conforming.

4.2 Capital Project Practice Benchmark Workflow Conformance Method

Measuring benchmark workflow conformance means comparing two workflows or comparing an event log to a workflow. A benchmark workflow is a workflow that is a standard or a target that an audited company attempts to pursue when implementing a process. The benchmark workflow typically is formed before the process is executed but can be modified after trial and error. When the process is executed, there are deliverables such as documents. If well-documented, document conformance measurement may be possible. However, documentation is often completed by humans; thus, it can have errors. Instead, if an event log recorded automatically by a computer is used for analyses, human errors may be reduced. Figure 4-4 is the functional requirements to achieve workflow conformance measurement.

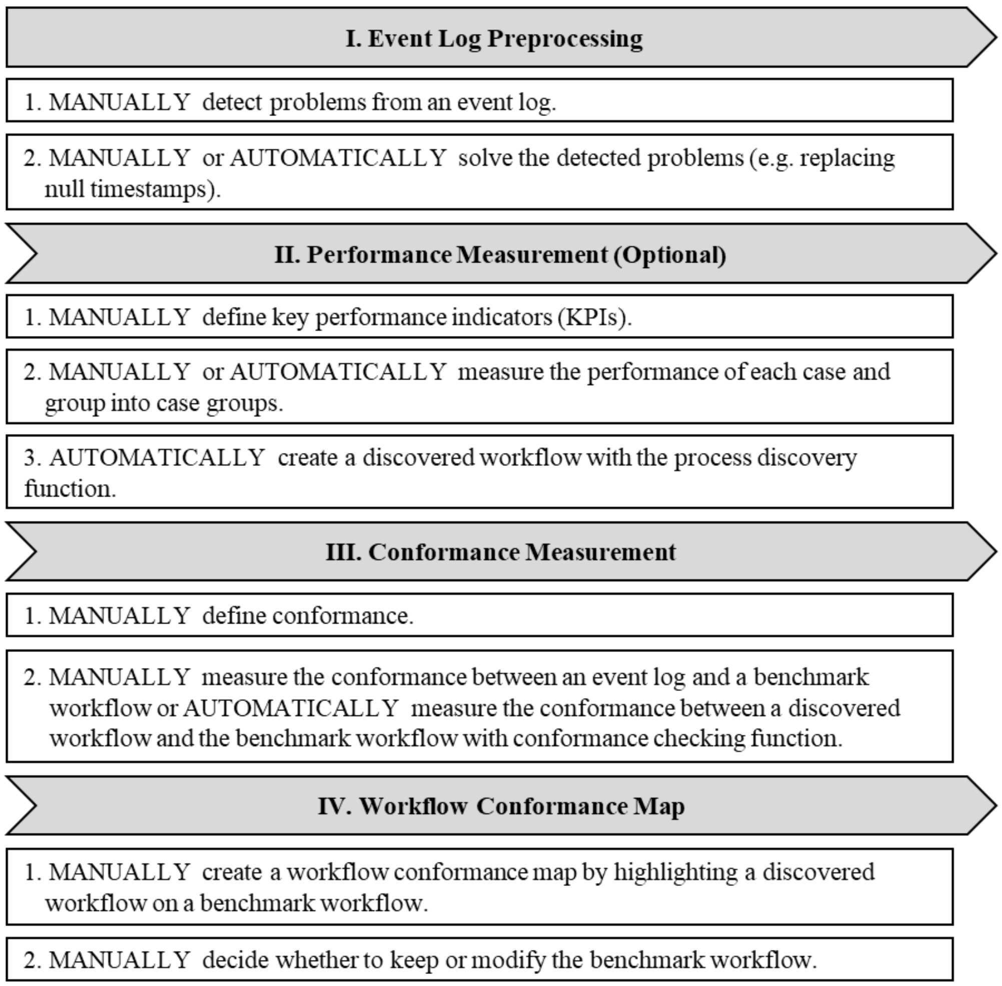


Figure 4-4: Functional Requirements of Workflow Conformance Measurement

The first module, event log preprocessing module (Module 1) is to clean the raw data into a format that can be understandable by both humans and machines. Problems must be detected and replaced properly within the context. This module is called the event log preprocessing module.

The second module is the performance measurement module (Module 2). In the performance measurement module, key performance indicators (KPIs) are defined. Then, performance for each case is measured. Next, each case is grouped into a case group according to its performance. A case group is a group of multiple cases with similar performances. Finally, a process discovery function creates an illustration of the discovered workflow from the selected case groups. A discovered workflow is a visual workflow derived from an event log. As opposed to a benchmark workflow which is a planned workflow, the discovered workflow presents an executed process. Whether it includes one or multiple cases, the process discovery function creates one workflow from start to end from a case group data. According to the grouping, there may be variations of discovered workflows.

The third module is the conformance measurement module (Module 3). In the conformance measurement module, conformance must be defined. Certain activities or connections may be skipped or inserted depending on the definitions. Then, a selected discovered workflow is compared to the benchmark workflow. Any non-conformances are detected. In the last module (Module 4), a workflow conformance map is created. It is an illustration to clearly visualize which path from the benchmark workflow the case groups have taken to achieve the goal.

4.2.1 Event Log Preprocessing Module

An event log is a sequentially recorded collection of activities. The event log can be created by humans or machines. If the event log is recorded in automated ways, the analyses are less likely to be biased. Companies attempt to implement new processes and enforce them to relevant departments to achieve better overall progress. When the departments implement these new processes, they may record the results manually. They may also record created and completed timestamps of each task or activity. However, these manual recordings by humans may not be accurate due to human errors or intentional deception. It also is tedious and doubles the work by increasing human time and effort. If these recordings are completed in automated ways, in the form of the event logs, many factors that incur bias and errors can be reduced. In response to the advance of information technology, automatic records of events have overcome the limitations of human recordings. However, there are some problems with automated event logs. Examples of problems and solutions are presented in Table 4-2 (Chen et al., 2019).

Table 4-2: Examples of Problems and Solutions for Event Log Preprocessing

No.	Activity		No.	Transition	
	Problem	**Solution**		**Problem**	**Solution**
❶	Duplicated names	Rename into different names	❺	Inconsistent format	Format to be consistent
❷	Inconsistent format	Format to be consistent (e.g., action + performer)	❻	Missing values due to instant occasion	Data imputation (e.g., completion time = start time)
❸	Ambiguous definitions	Remove ambiguity	❼	Missing values due to overtime or group recipients	Data imputation
❹	Missing information	Add information	❽	Missing values in abandoned activities	Ignore the missing values

There are two types of problems incurred by activities or timestamps. Depending on the context, each problem must be resolved differently. The reason for these problems is that software engineers who initially format the event logs and auditors who use the event logs for analysis do not share the same knowledge or values. Due to the lack of understanding, knowledge, or experience in these respective domains, companies may not have documented the event logs most effectively. With the poorly structured event logs, the software may not be able to provide accurate data. Therefore, there is a need to preprocess event logs to improve their quality. The preprocessing module is essential not only for clarifications but also to enable the logs to be run by the process mining software packages.

Figure 4-5 represents example problems that are mentioned in Table 4-2 with corresponding numbers. There may be more than one solution to each problem. Therefore, it is critical to replace or impute data manually depending on contexts and circumstances. However, this may be completed with caution since information can be lost or manipulated. If so, the results cannot be an accurate representation. Because some process mining software cannot process blanks or nulls which are timestamp problems, blanks or nulls need to be replaced.

WF_ID	activitydisplayname	createddatetime	OwnershipDateTime	completeddatetime	ResponseBy	Name	CurrentStatus
216	Verify Details	3/30/11 17:59	31-03-2011 13:36	NULL	NULL	Amin	Closed
216	Verify Details	3/30/11 17:59	31-03-2011 13:36	NULL	NULL	Blake	Closed
216	Verify Details	3/30/11 23:59	31-03-2011 13:36	3/31/11 13:36	NULL	Amin	Send On
216	Change Request Participants Verification	3/31/11 1[illegible]	31-03-2011 19:52	3/31/11 19:53	NULL	Amin	Send On
[illegible]	Review (Engineer)	3/31/11 19:53	01-04-2011 17:21	4/4/11 22:45	05-04-2011 23:00	Tracy	Send for Review
216	Change Request Participants Verification	4/4/11 22:45	04-04-2011 22:45	4/5/11 19:06	NULL	Tracy	Send On
216	Review (Participants)	4/5/11 19:06	07-04-2011 16:15	4/7/11 16:19	11-04-2011 22:00	Vic	Send On
[illegible]	Review (Participants)	4/5/11 19:06	05-04-2011 19:07	4/5/11 20:34	11-04-2011 22:00	Karen	Send On
216	Review (Participants)	4/5/11 19:06	06-04-2011 12:54	4/6/11 13:08	11-04-2011 22:00	Kirk	Send On
[illegible]	Review (Participants)	4/5/11 19:06	05-04-2011 23:13	4/5/11 23:16	11-04-2011 22:00	Andrew	Send On
[illegible]	Approve (Engineer)	4/7/11 16:19	09-04-2011 [illegible]	NULL	09-04-2011 19:28	Tracy	Closed
[illegible]	Approve (Engineer) Warning	4/9/11 19:28	14-04-2011 21:11	4/15/11 21:18	NULL	Tracy	Reject
[illegible]	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Randy	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Tim	Information
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Vic	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Karen	Information
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Andrew	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Kirk	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Duffy	Information
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Tracy	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Kevin	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Jack	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Don	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Amy	Deleted
216	Rejected Notification	4/15/11 21:18	NULL	NULL	NULL	Michael	Deleted
216	Rejected Close Out	4/15/11 21:18	15-04-2011 22:41	4/15/11 22:42	NULL	Amin	Send On
241	Verify Details	4/5/11 15:47	06-04-2011 13:26	NULL	NULL	Faisal	Closed
241	Verify Details	4/5/11 21:47	06-04-2011 13:26	4/6/11 13:27	NULL	Amin	Send On

Figure 4-5: A Spreadsheet Example of Event Log Problems before Preprocessing

In general, if well communicated, activity names are not an issue. The activity names can be confusing when they are not clearly defined. The example (Figure 4-5) is unclear because they are too simple and do not entail all the necessary information. Hence, if the auditor is not involved when activities are defined, it is important to align the implicated definitions before additional analyses. Since an event log is often a by-product created for backup to data loss accidents, it is not always the main interest. That is why some activity names are overlapping and some may even not indicate what it means. Some naming may not include necessary information. For instance, because "Change Request Participants Verification" is repeated (❶), it may be considered the same activity. However, they are two distinctive activities. For other cases (❷❸❹), their activity names are unclear due to the format - "activity (participant)." The roles of participants are unclear (e.g., some are active participants and others are participants who are affected by the activities).

Timestamp issues are more critical because process mining software may not further proceed if these issues are unresolved. Timestamps indicate the created time and the completed time. There are two main problems with timestamps. One is the inconsistent formatting, and the other one is the null timestamps. The null timestamps occur due to many reasons. For example, some activities happen instantly (❻), some are aborted, or some are not completed before deadlines (❼). These timestamps need to be replaced. The reason for difficulties in automation for the preprocessing module is that each case needs to be resolved independently based on the context. For example, for the null case of an activity's completed timestamp, sometimes the previous activity's completed time is more suitable whereas other times the current activity's created time makes more sense.

4.2.2 Performance Measurement Module

Performance measurement is optional but is related to conformance in the following section. When measured, the benchmark workflow can be improved. Before measuring performance, key performance indicators (KPIs) must be defined. The forms of performance may differ depending on the values and objectives of a company. It can be cost savings, duration reductions, or improved customer satisfaction. After KPIs are defined, the performance of each case, which is identified by an ID, needs to be measured to group the cases. Case ID is one of the three necessary components in process mining (the others: activity, timestamp). Each case has a start and an end. Within this interval, there are activities. Each case may have a label indicating its performance. If there is more than one KPI, there can be multiple labels. However, each case can only gain one rating for the respective KPIs. If it is a pass-or-fail test, either pass or fail is given to a case, not both.

Once rated, groups can be formed according to the ratings. After grouping, using the process discovery function, a discovered workflow can be created. Since there is no consensus on the terminology, in this thesis, it (called discovered model or process by some software) is called a discovered workflow. A discovered workflow is defined as a path created by a process discovery function to visually illustrate activities and transitions that occurred through a process start to end. A discovered workflow can contain single or multiple cases. Process discovery function constructs a representation of the current processes of an audited company automatically. In other words, the process discovery function converts an event log into a visual workflow form. There are pros and cons to this conversion because by generalizing the data, some data may be lost. For conformance measurement, process discovery is not necessary. When raw event log data is compared against the benchmark workflow, the results may be more accurate. This can be completed manually.

4.2.3 Conformance Measurement Module

In the performance measurement module, performance is defined and measured. In the conformance measurement module, conformance is defined and measured. The conformance can be defined with conformed activities and transitions. That is, if the activities or transitions that appear in the discovered workflow exist in the benchmark workflow, it is in conformance. If the activities or transitions that appear in the discovered workflow do not exist in the benchmark workflow, it is not in conformance. In the conformance measurement module, a discovered workflow is compared to the benchmark workflow. The benchmark workflow is sometimes called the target model or process model. Comparing event logs or discovered workflows against the benchmark workflow is generally known as conformance checking in the process mining domain. Conformance can be measured manually or with the process mining software packages. Depending on how conformance is defined and how the software is programmed, the results vary. In process mining software packages, a conformance checking function compares benchmark workflow with a discovered workflow or an event log itself. It is used to check if the actual implementation of a process conforms to the benchmark workflow.

There are two ways of checking conformance in the process mining field, namely, a model-to-model and log-to-model comparison. A model-to-model comparison refers to comparing a discovered workflow against a benchmark workflow. Because both benchmark workflow and discovered workflow are in the form of workflow, which is a two-dimensional diagram, they both are referred to as models. On the other hand, an event log is a one-dimensional list of records. When converting a log into a discovered workflow, there may exist information losses.

There are benefits to both comparisons. A model-to-model comparison provides visualizations that ease comprehension while a log-to-model comparison involves less information loss. Because an event log is a record of numerous arrays of characters and values, the model-to-model comparison is helpful to simplify the complexities of the process. However, when it is too simplified, it may not represent the original event log. When it is too generalized, it may not be precise enough. A log-to-model comparison refers to comparing an event log against a benchmark workflow. The event log entails specifics such as created and completed time. It is unfiltered and can provide additional information. However, because it is too specific, it is difficult to generalize. Because it is not visually friendly, it takes more time and effort to detect non-conformance.

There are a few ways to define conformance. For example, footprint-based conformance checking allows the model-to-model comparison as well as the log-to-model comparison (Rozinat, 2005). A footprint is a combination of direct sequences from an activity to another activity. By comparing a footprint of a discovered workflow against a benchmark workflow or an event log against the benchmark workflow, the matching footprint is defined as conformance. However, since this method does not account for the frequencies of occurrence, it can be biased. Whether an activity happened once or a thousand times, it just presents as it has occurred. Some examples follow for understanding conformance measurement. Figure 4-6, two 2-by-2 matrices (footprints) are created based on a benchmark workflow and an event log, respectively.

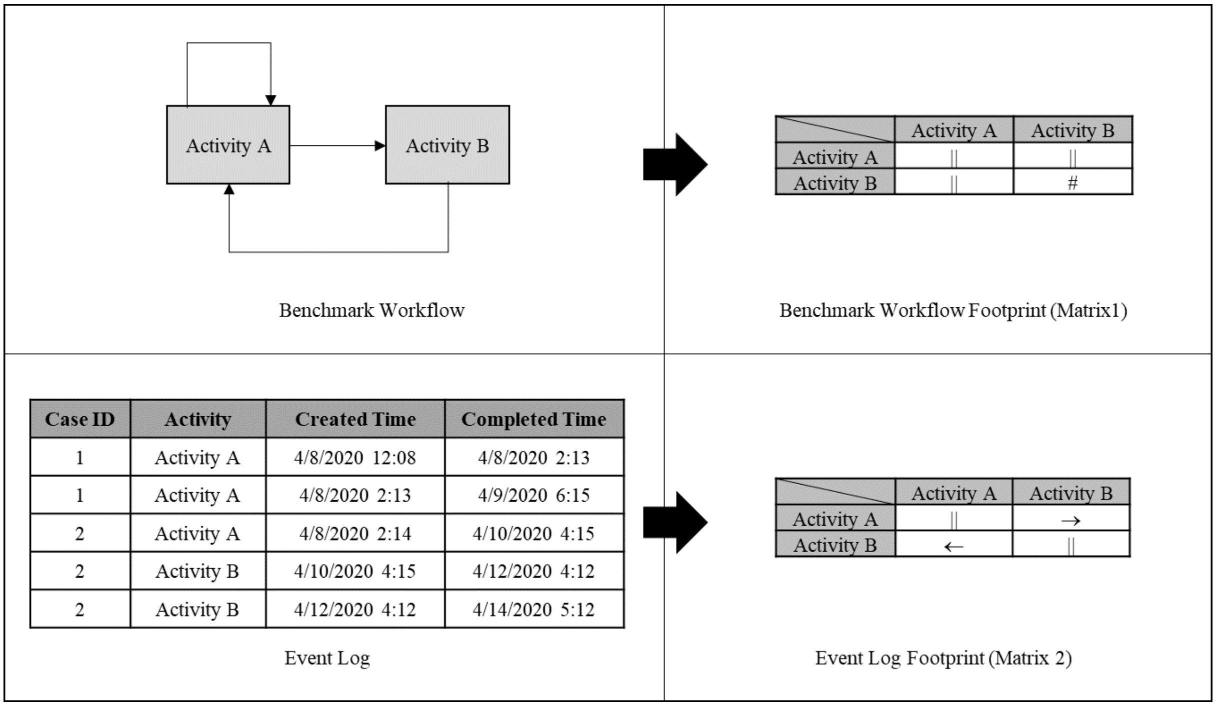


Figure 4-6: Footprint-based Conformance Checking Method

Figure 4-6 can be interpreted as if Activity A is directly followed by Activity A, || is assigned in the AA cell. The AB cell and BA cells are labeled differently in two matrices as the benchmark workflow matrix has a parallel symbol (||), while the event log matrix has an arrow symbol (←/→). The BB cell is also different since in the benchmark workflow, Activity B cannot be followed by Activity B (#), whereas in the event log, Activity B is sequential (||). With these footprints and equation (3), conformance can be derived in numeric form. Thus, it is 25% (1 match/4 cells) conformance.

Table 4-3: A Footprint-based Conformance Checking Representations and Definitions

Representation	Definition
A→B	There is at least one case that A is directly followed by B
A←B	There is at least one case that B is directly followed by A
A‖B	There is at least one case that A is directly followed by B and B is directly followed by A
A#B	There is no case that A is directly followed by B or B is directly followed by A

The conformance may be derived by Equation (3).

$$Conformance = 1 - nonconformance$$

$$= \frac{Total\ number\ of\ matching\ cells}{Total\ number\ of\ cells} \quad \dots\dots\dots\dots\dots\dots\dots(3)$$

Another conformance checking method that is used in some software is checking the existence of activities. To derive the conformance, the ratio of the number of inserted and skipped activities over the total number of activities is computed. The definitions are provided in Table 4-4.

Table 4-4: Terms and Definitions for Measuring Conformance

Terms	Definitions
Conformance	1 – nonconformance
Nonconformance	(number of inserted activities + number of skipped activities) / total number of activities
Inserted activities	activities that exist in the discovered workflow that do not exist in the benchmark workflow
Skipped activities	activities that do not exist in the discovered workflow that do exist in the benchmark workflow
Conforming activities	activities that do exist both in the discovered workflow and the benchmark workflow
Total activities	inserted activities + skipped activities + conformed activities

Another way to check conformance is to identify types of activities and transitions. Checking whether they exist or not may provide meaningful information. There are other ways to define conformance. Once it is defined, the conformance may be measured manually or with software packages. Then, it must be reported to the workflow conformance map module.

4.2.4 Workflow Conformance Map Module

A workflow conformance map is a visualization of a comparison between a benchmark workflow and a discovered workflow. A benchmark workflow is a collection of all possible paths that are expected when the process occurs. A discovered workflow may include a single case or multiple cases. After the performance and conformance measurement modules, a workflow conformance map can be created. The workflow conformance map contains the benchmark workflow as a backdrop and marks the conformed activities and transitions of the discovered workflow over the backdrop. A Swimlane flowchart to derive the conformance map is presented in Figure 4-7. Gray boxes indicate potential semi-automation.

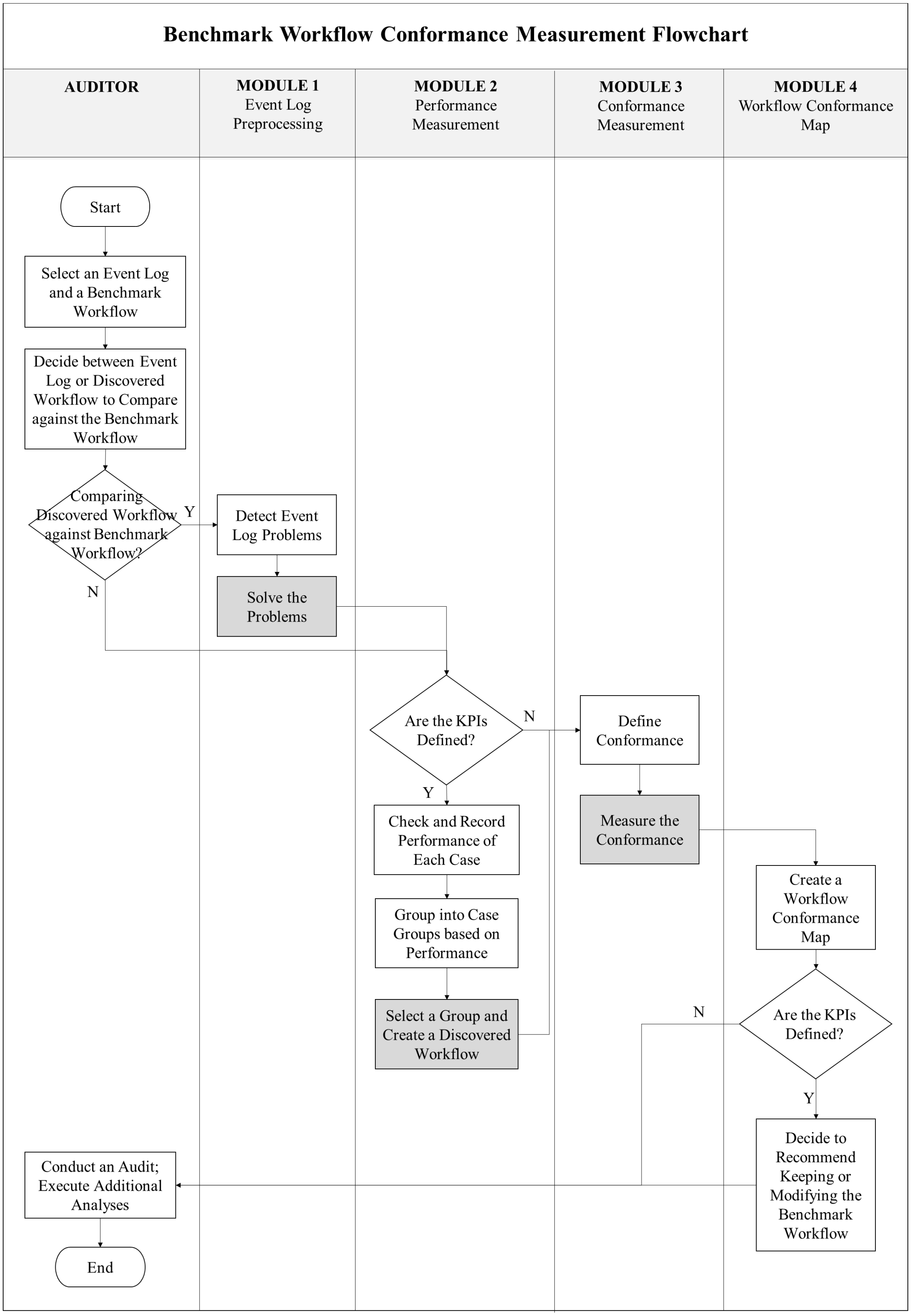


Figure 4-7: Swimlane Flowchart for Workflow Conformance Measurement

To create a workflow conformance map, first, an auditor may select an event log and a benchmark workflow to compare with. Then, the auditor must decide whether to compare the raw event log data or a processed discovered workflow derived from the event log against the benchmark workflow. When the event log is converted into the discovered workflow, the conversion requires the event log preprocessing module (Module 1) due to the software limitations. When the event log is compared to as-is, because the manual process takes place, no conversion is required.

Next, key performance indicators (KPIs) need to be defined if the auditor intends to investigate thoroughly and provide recommendations for improvement in the existing benchmark workflow. This step (in Module 2) is not mandatory when conformance solely needs to be measured. When KPIs are defined, for every case (with different ID), performance needs to be measured. Then the auditor can group the cases into case groups with similar performances. With this data, the auditor can analyze further by finding patterns within and across the cases. The selected case group may create a discovered workflow with the process discovery function provided by software packages.

Once the performance measurement module is completed, conformance must be defined and measured (Module 3). Conformance measurement can be completed through the conformance checking function provided by some software packages. Based on the results from either manual or software packages, the auditor must record by drawing a path on the benchmark workflow (Module 4). The drawing must include marks if there are non-conformances. With this information, the auditor can further analyze the process and the workflow.

Chapter 5 Framework Analysis and Toolset Verification for Assisting Semi-Automated Practice Conformance Measurement

In order to semi-automate the framework, appropriate toolsets are required. No standard toolset is designed for the purpose of document or workflow conformance measurement while other fields have representative software like Matlab for machine vision or SPSS for statistics. Software packages were selected to create toolsets that were accurate enough to suit the purpose of each module of measuring practice conformance. Companies can utilize this framework and integrate their existing specialty software or other variations by building or customizing their own toolsets.

Document conformance measurement and workflow conformance measurement complement each other to achieve practice conformance measurement. Whereas document conformance measurement can be used to check the states of companies, workflow conformance measurement can be used to check the control logics, a key part that controls the operations. Corporate documents are created by humans; thus, may have human errors. Event logs are created automatically; yet may have software errors. Co-existence of document and workflow conformance measurement method allows comprehensive analysis.

Given commercial technologies and availability of data in machine-readable forms, semi-automation for practice conformance measurement is perceived possible in this study. The objective of this chapter is to verify the applicability of the commercial software packages when they are used to meet functional requirements addressed by this study. Functional analysis is performed to determine whether each software reaches adequate accuracy. Adequate accuracy is the degree of accuracy that can be accepted and used. Twelve commercial software packages for document conformance and two commercial software packages for workflow conformance are tested and evaluated. For each functional requirement, the most suitable software is selected. With selected commercial software packages, pilot studies are conducted to perform functional demonstrations.

5.1 Semi-Automated Practice Guideline Document Conformance Measurement Method and Toolsets

Text extraction, keyword extraction, and text matching tools are introduced and evaluated for verification. Commercial software packages need to reach adequate accuracy for the specific purposes of this study. For example, although plagiarism detection tools or website content duplication detection tools may be applied to text matching purpose, they may not be sufficient since their original intentions are unlike the intention of semi-automating conformance measurement. In other words, a tool may be suitable for its original purpose but not for the purpose of this study. Compared to the workflow conformance measurement toolsets, a wide variety of software packages of document conformance measurement toolsets are investigated. For each module, some software packages are selected to compare costs, durations, and accuracy.

5.1.1 Analysis of Text Extraction and Toolset Verification

According to Figure 4-2, the text extraction module can be characterized as semi-automated. Texts either appear as a handwritten or typed form. When it comes to a typed form, there are read-only formats such as Adobe Portable Document Format (.pdf) and modifiable formats such as Microsoft Word (.docx). Some companies prefer to store documents in PDF files while others prefer Word documents because they entail different purposes.

A corporate document can potentially be written in a Word document and be kept in the same format. However, practice guideline documents are often distributed as PDF files, since they are intended for viewing only, and they are kept with fewer data storage. Yet, the PDF files provide an electronic image of texts and graphics. Therefore, there are some problems to be addressed when PDF files need to be reverse translated into a machine-readable or modifiable document. An example of limitations of text extraction software packages follows. When a word is separated by a new line, a Word document sometimes connects with a hyphen (-) automatically. For instance, a 'word' is divided into 'wor-' and 'd' in a Word document; then, the Word document is converted into a PDF file where 'wor-' and 'd' are saved as images. When the PDF file is reconverted into a Word document, rather than returning 'word' it prints 'wor-d' which is not the word that is originally intended. Such examples are limitations of conversions, and they represent information losses. Though text extraction tools are necessary for some situations, it is recommended to skip this module if possible. The text extraction tools can be divided into two types: (1) PDF file conversion to Word, and (2) hand-written documents to machine-readable documents.

For PDF conversion, a typed document but saved as a PDF format is the scope of part (1) above. How a hand-written document saved in PDF file format is handled is covered as part (2) from above. There are problems such as fonts that do not exist in software packages for conversion from PDF files to Word documents. However, since type-set letters are at least consistent, accuracy for (1) PDF conversion is higher compared to (2) hand-written document conversion. Table 5-1 summarizes how the software packages are evaluated.

Table 5-1: Legend for Evaluation Criteria When Converting File Formats

Rating of Accuracy	Definition
A	Correctly opens and returns the texts without any error
B	Correctly opens with minor errors including few misrecognized letters
C	Opens with major errors and not readable
D	Errors exist in opening a file

Rating A is assigned to software packages that correctly open and return the texts without any issue. The fonts may differ from original files. Rating B is assigned to software packages that have minor errors such as the letter "t" misrecognized as the letter "f," or the letter "I" as the letter "l." Hyphenated words that are in the original document which are not recognized as one word are considered a minor error. Rating C is assigned to software packages that have major errors in which the document is not readable although the software packages can open the files. Rating D is assigned when the software packages cannot open the files.

Table 5-2 represents the comparisons of file conversion software packages. Cost, duration, and accuracy (ratings) of the conversion results are presented. For the verification, a 90-page PDF document is converted to a Word file. Cost refers to the initial cost of software packages. Duration refers to the duration of the conversion process. Software package installation time is excluded. Labor costs are excluded. Depending on the situation, an audited company or auditor may already have these software packages.

Table 5-2: Verification Results for PDF File Format Conversions

Criteria	Toolset 1 Google OCR	Toolset 2 Word 365	Toolset 3 Adobe Pro	Toolset 4 WPS
Cost	$0	$70/yr	$180/yr	$30/yr
Duration	18 sec	42 sec	26 sec	10 sec
Accuracy (Rating)	B	B	A	A

For the best accuracy and reasonable price, Toolset 4 (WPS) is selected for the pilot study and case studies. The document, that is used for the software verification test, has some images and is a double column document. Thus, there are some difficulties when this PDF file is converted into Word format. For example, Toolset 1 (Google OCR) and Toolset 2 (Microsoft Word 365 for Business) were not able to recognize double column texts; thus, returning the results that require manual re-formatting work.

Table 5-3 represents the comparisons among hand-written document conversion software packages. Cost, duration, and accuracy (ratings) of the conversion results are presented. For the verification, a one-page cursive and print handwriting document image is converted to Word file. Cost refers to the initial cost of software packages which are the same as Table 3-4. Duration refers to the duration of the conversion process. Software package installation time is excluded. Labor costs are excluded. Depending on the situation, an audited company or auditor may already have these software packages.

Table 5-3: Verification Results for Hand-Written Document Conversions

Criteria	**Toolset 1 Google OCR**	**Toolset 2 Word 365**	**Toolset 3 Adobe Pro**	**Toolset 4 WPS**
Cost	$0	$70/yr	$180/yr	$30/yr
Duration	16 sec	N/A	15 sec	10 sec
Accuracy (Rating)	B	N/A	D	D

For the verification, a scanned hand-written document is converted to machine-readable texts. The test results are presented in Appendix A. This function is perceived useful but is not a priority for the pilot/case studies since documents are rarely handwritten nowadays. Accuracy is not enough either. However, Toolset 1 (Google OCR) provided relatively more accurate results. This may be due to the large volume of trained data that is required for machine learning in optical character recognition (OCR).

Most companies use machines to record documents. However, in some cases, it is easier and faster to deal with paper documents, especially in the construction industry. Checklists or forms are some of the examples that are recorded manually. Typically, in these cases, scanned documents are then saved in the company database. With image processing technologies, information from hand-written documents can be directly extracted to the system in automated ways.

5.1.2 Analysis of Information Type Detection

The information type detection module is an optional module depending on the definition of the conformance. It may increase or decrease the accuracy according to how the module is applied. For instance, there are many charts in alcohol and drug-related documents. Due to time constraints, not every piece of document can be examined. By acknowledging the importance of drug concentration limit charts, comparison between two charts may provide meaningful insights in the shorter term.

Another example is management plans. These differ from company to company and oftentimes may not include the same contents suggested by institutions. However, the types of information may overlap. That is, the "schedule/timeline" element may exist in both documents though they are not alike. By solely detecting the existence of the information types, conformance can be measured (as conformance is achieved). However, the time that is required for this module varies due to the manual tasks; thus, it is suggested to be applied with consideration. With current technology, it is difficult to specifically distinguish the type of information in automated ways.

5.1.3 Analysis of Keyword Extraction and Toolset Verification

The keyword extraction module is also an optional module. It may be used in different orders across other modules. For instance, after the text matching module, the keyword extraction module may be used for further comparisons of the non-matching portions derived by the text matching module. The keyword extraction module may be used before the text matching module to use the keywords as indices for sampling when there are time constraints.

As the keyword extraction toolset, word cloud generators are examined. Word cloud generators create a shape that composes a collection of words. The graphic is based on the frequencies of words from the input. When the frequencies are higher, the sizes of the words are larger. Though they seem simple to use, their basic operating concepts are founded upon natural language processing (NLP). Commonly used words, also known as stop words, are filtered through the system because their existences are meaningless to users (e.g., is, a, the, in, and, not). Depending on how the software is programmed, some stop words (e.g., will) do appear in some of the word cloud generators.

Some software packages involve stemming and lemmatizing in their word cloud generator. Stemming and lemmatizing steps both retain the base meanings of words but remove the last few characters. Stemming and lemmatizing differ in that stems are often grammatically incorrect whereas lemmas are the base forms of words that lead to grammatically correct words. For example, “manage” has various forms such as “management,” “managing,” and “managed.” When these words are stemmed, the results will be “manag” whereas “manage” when lemmatized. These steps are to retain semantically meaningful parts and remove unnecessary parts so that computers can recognize the forms of the same meaning words as the same. This functionality is helpful to reduce the list of words that appear in the results. Advanced word cloud generators use other keyword extraction methodologies to recognize collocations and co-occurrences.

Costs, duration, and the accuracy of the word cloud generator are compared. Cost includes software package cost but not labor cost. Duration includes the processing time of the software packages but not the preparation time. When it comes to accuracy, it is differentiated into two criteria: (1) stop word recognition and (2) stemming and lemmatization (Table 5-4).

Table 5-4: Legend for Evaluation Criteria When Converting File Formats

Rating of Accuracy	Definition	
	Stop Word Recognition	Stemming and Lemmatization
A	All stop words are detected	Successfully implemented
B	Some stop words are detected	Partially implemented
C	No stop word is detected	Not implemented

For (1) stop word recognition, if all stop words are excluded from the list, the toolsets are rated A. If the stop words are not detected correctly, the toolsets are rated B. If the stop words are not detected at all, the toolsets are rated C. For (2) stemming and lemmatization, if stemming and lemmatizing have been successfully implemented, the toolsets are rated A. If stemming and lemmatizing have been partially implemented, the toolsets are rated B. if stemming and lemmatizing have not been implemented, the toolsets are rated C.

A one-page document is tested with four types of word cloud generator software packages. The results are summarized in Table 5-5. Cost, duration, and accuracy (ratings) of the conversion results are presented. For the verification, a one-page interface management document is used to create a word cloud. Cost refers to the initial cost of software packages. Duration refers to the duration of the conversion process.

Table 5-5: Verification Results for Keyword Extraction Toolsets (Word Cloud Generators)

Criteria	Toolset 5 WordClouds	Toolset 6 WordArt	Toolset 7 TagCrowd	Toolset 8 WordItOut
Cost	$0	$0	$0	$0
Duration	4 sec	5 sec	2 sec	1 sec
Accuracy (Stop Words)	B	A	A	B
Accuracy (Stemming)	C	B	B	C
Ease of Use	PDF accepted	PDF not accepted	PDF accepted	PDF not accepted

All software packages provide the service to view the word cloud and download the list in an excel spreadsheet with no cost. Ease of use criteria is included to provide information on whether the PDF file can be processed or not. This function is perceived vital since most documents are in PDF format. Appendix B provides graphical results of the keyword extraction toolset verification test.

5.1.4 Analysis of Text Matching and Toolset Verification

The text matching module is also optional, but some potential appears to exist. To detect the correlations among documents whether they are related, matching, or conforming to each other, the text matching module and toolsets are practical. Plagiarism detection tools or website content duplication detection tools are some of the candidate toolsets that can be applied for the semi-automated conformance measurement purpose. The plagiarism detection tools are used typically in an academic setting to recognize similarities in language, thoughts, ideas, or expressions. One of two types of plagiarism detection tools (i.e., text comparison tools) may be used for text matching. The other type of tool searches against multiple sources from the web (e.g., iThenticate, Turnitin). The website content duplication detection tools also apply the technologies from the second type of plagiarism detection tools. Though plagiarism must be detected throughout all sources when it comes to intellectual properties, for the purpose of the conformance measurement of multiple documents, text comparison technology is perceived to be suitable. Table 5-6 summarizes the criteria for accurate text matching software packages.

Table 5-6: Legend for Accuracy Criteria for Text Matching Tools (Plagiarism Detection)

Rating of Accuracy	Definition
A	Successfully identifies most matching including some synonyms (dictionary- or context-based)
B	Successfully identifies most matching; but not synonyms (dictionary- or context-based)
C	Unsuccessful in identifying most matching

If most matching texts are successfully identified by the software packages, the software packages are rated A. These software packages must recognize some synonyms as matching. If most matching texts are successfully identified, but, synonyms are not recognized as matching, the software packages are rated B. If software packages are unsuccessful in identifying most matching, they are rated C. Costs per use (page) are compared for the cost excluding labor costs. Duration includes the processing time of the software packages but not preparation time. Ease of use criteria is included to provide information on whether PDF file can be processed or not. This function is perceived vital since most documents are in PDF format.

Four software packages are verified with two documents related to alcohol and drugs. The results are summarized in Table 5-7. The original documents are presented in Appendix C. Cost, duration, and accuracy (ratings) of the conversion results are presented. For the verification, two documents of two-page text files are entered into the software packages. Cost-per-use refers to cost-per-page to compute the data. Duration refers to the duration of the computation process.

Table 5-7: Verification Results for Text Matching Toolsets

Criteria	Toolset 9 Copyleaks	Toolset 10 Copyscape	Toolset 11 Plagscan	Toolset 12 Countwordsfree
Cost-per-Use	$0.05/pg	$0	$0.23/pg	$0
Duration	22 sec	1 sec	53 sec	2 sec
Accuracy	A	B	B	B
Ease of Use	PDF accepted	PDF not accepted	PDF accepted	PDF not accepted

Alcohol and drug-related documents are likely to be guidelines or policies. Thus, they are expected to have some matchings regardless of their sources. Durations for processing the documents are less than one minute for all software packages. Mostly, the software packages that allows PDF file attachment requires cost-per-use. Toolset 9 (Copyleaks) is selected for further studies due to its accuracy and ease of use.

The results of the text matching module are presented in the document conformance map module (Section 5.1.5). The document conformance maps are automatically created in the text matching module with the software packages.

5.1.5 Analysis of Document Conformance Map

A document conformance map is an illustration of conformances and non-conformances that visually guides its users mainly for initial decision making such as sampling or for further analysis. Once the modules that compose the method are decided, corresponding conformance maps can be created. Depending on which modules are selected, the levels of detail may be altered. For instance, if the information type detection module is solely selected, the conformance map may demonstrate the existence of elements but not the correctness of the elements. If text matching is followed by the keyword extraction module, the conformance map may elaborate in detail. For example, even though text matching has not been achieved, documents may be conforming since keywords that are extracted match with each other.

Two of the automatically created text matching results (i.e., Toolset 9 (Copyleaks) and Toolset 10 (Copyscape)) are presented in Table 5-8 and Table 5-9. Identical portions are marked yellow, minor changes are marked in pink, and related meanings or synonyms (dictionary- or context-based) are marked in gray. Two software packages return similar results; yet, Toolset 9 (Copyleaks) is selected since it catches minor changes and synonyms. Another difference between Toolset 9 (Copyleaks) and Toolset 10 (Copyscape) is that Toolset 9 (Copyleaks) distinguishes two comparing documents into original and suspect, whereas Toolset 10 (Copyscape) does not. However, they both provide two types of matching rates. These matching rates indicate the overall similarities between the two documents. Toolset 9 (Copyleaks) has strength in that more than two corporate documents (suspects) can be compared at once against a practice guideline (original). The results of other software packages are attached in Appendix D.

Table 5-8: Document Conformance Map Created by Toolset 9 (Copyleaks)

Practice Guideline	Corporate Document
4.0 SOCIAL SITUATIONS 4.1 Company Sponsored Social Events In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community. Responsible alcohol use is permitted at Company sponsored social functions held away from Company premises, which must have the prior approval of a Vice President level or above, and will be conducted in accordance with the Company's hosting guidelines. Anyone who attends and consumes alcohol must not be returning to or going to work after the event. 4.2 Business Hosting Consistent with the above, if alcohol is made available to guests in the course of conducting business (e.g., restaurant meeting), employees are expected to use judgment and be responsible in hosting others. 5.0 CONSEQUENCES OF A POLICY VIOLATION 5.1 General Requirements Any violation of the provisions of this Policy may result in corrective action or termination of employment. Management has the authority and discretion to hold out of service any individual who is believed to be involved in an incident that could lead to corrective action pending the results of the investigation. The appropriate action in a particular case depends on the nature of the policy violation and the circumstances surrounding the situation; the severity of the violation may warrant entering the corrective action process at different levels or termination of employment. For all employees any confirmed situation of drug trafficking on Company premises will result in termination of employment. A positive drug test, or an alcohol test result of .04 BAC or higher, or a refusal to test are all considered a violation of this Policy. An alcohol test of .02 BAC, or higher, for anyone working at a designated dry site is considered a violation of this Policy. In all other situations, an employee who has an alcohol test result of .02 to .039 BAC in a reasonable cause or post incident situation will be removed from duty until considered safe to return (at a minimum not before	8. SOCIAL SITUATIONS In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community. Subject to any site specific limitations, responsible Alcohol use may be permitted at Company sponsored social functions with appropriate prior approval. Alternative transportation arrangements will be made available when possible. Consistent with the above, if Alcohol is made available to Company guests in the course of conducting Company Business (e.g.,, restaurant meetings), Employees are expected to use reasonable judgment and be responsible in hosting others, and remain in compliance with the Policy and Supporting Standards. Procedures for hosting events are set out in the Social and Business Hosting Standard. 9. CONSEQUENCES OF A POLICY VIOLATION a) General Requirements: Any violation of this Policy and Supporting Standards may result in discipline up to and including termination of employment. In all situations, an investigation will be conducted to verify that a Policy or Standard violation has occurred. The appropriate discipline in a particular case depends on the nature of the Policy or Standard violation and the circumstances surrounding the situation. The severity of the violation will warrant entering the discipline process at different levels. General violations of this Policy include: i. failure to comply with the Policy and Supporting Standards; ii. a positive Alcohol or Drug test (refer to the Alcohol and Drug Testing Standard); or iii. a Failure to Test. b) Referral for Assessment: After any confirmed positive Alcohol and Drug test, an Employee may be referred by Health and Wellness to a Substance

their next work day or shift) and may be subject to corrective action. After any confirmed violation, the employee may be referred for a SAE assessment to determine whether there is a need for a structured treatment program. 5.2 Conditions for Continued Employment Should the Company determine that employment will be continued in a specific circumstance following a policy violation, the employee will be required to enter into an agreement governing their continued employment which may require any or all of the following actions, or any other condition appropriate to the situation: • adherence to any recommended treatment and aftercare program; • maintenance of sobriety and satisfactory performance on return to duty; • successful completion of a fit for duty test; • ongoing unannounced testing for a period determined on a case by case basis taking into account the recommendations of the SAE; • adherence to any rehabilitation conditions or requirements; and • no further violations of the Policy during the monitoring period. Failure to meet the requirements of the agreement during the monitoring period will be grounds for corrective action up to and including termination of employment as set out in the agreement.	Abuse Professional for a Substance Abuse Assessment (refer to the Substance Abuse Assessment Standard). Failing to meet with the Substance Abuse Professional or attend a scheduled Substance Abuse Assessment is a violation of this Policy. c) Conditions for Continued Employment: Should the Company determine that employment will be continued after a violation of the Policy or Supporting Standards, the Employee will be required to enter into an agreement governing their continued employment which may require any or all of the following actions, or any other condition appropriate to the situation: i. temporary removal from their position; ii. adherence to any recommended treatment and aftercare program; iii. successful completion of a return to work Alcohol and Drug test; iv. ongoing unannounced follow-up Alcohol and Drug testing for the duration of their agreement; v. adherence to any ongoing rehabilitation conditions or requirements; and vi. no further Policy or Standard violations during the monitoring period. Failure to meet the requirements of the agreement will be grounds for discipline up to and including termination.

Toolset 9 (Copyleaks) is capable of detecting minor changes and related meaning. Implications are elaborated in Table 5-9.

Table 5-9: Implication/Analysis of Corporate Document Compared to Practice Guideline

Corporate Document	Changes (Identical/ Added/ Removed/ Modified)	Implication/Analysis
In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community.	Identical	General statement. Identical match implies that practice guideline has been referenced
Responsible alcohol use is permitted at Company sponsored social functions held away from Company premises, which must have the prior approval of a Vice President level or above, and will be conducted in accordance with the Company's hosting guidelines. Anyone who attends and consumes alcohol must not be returning to or going to work after the event.	Removed	The intention of the removed statement must be investigated. It may be due to customization (e.g., since the company do not have hosting guidelines) or it may be unintentional negligence
Subject to any site specific limitations, responsible Alcohol use may be permitted at Company sponsored social functions with appropriate prior approval. Alternative transportation arrangements will be made available when possible.	Added	The additional statement must be carefully investigated since it is intentional customization. The contents should align with the practice guideline including added portions.
Consistent with the above, if Alcohol is made available to Company guests in the course of conducting Company Business (e.g.,, restaurant meetings), Employees are expected to use reasonable judgment and be responsible in hosting others, and remain in compliance with the Policy and Supporting Standards.	Modified	The minor changes implies that the practice guideline has not been randomly copied.
The appropriate discipline in a particular case depends on the nature of the Policy or Standard violation and the circumstances surrounding the situation. The severity of the violation will warrant entering the discipline process at different levels.	Modified	The term “action” and “may” changed into “discipline” and “will.” Though they are synonyms, different words have different nuance and different level of enforcement.
General violations of this Policy include: i. failure to comply with the Policy and Supporting Standards; ii. a positive Alcohol or Drug test (refer to the Alcohol and Drug Testing Standard); or iii. a Failure to Test.	Added	In fact, it is modified and not added; Corresponds to “A positive drug test, or an alcohol test result of .04 BAC or higher, or a refusal to test are all considered a violation of this Policy.” This is not detected by the toolset since the wording is significantly different
b) Referral for Assessment: After any confirmed positive Alcohol and Drug test, an Employee may be referred by Health and Wellness to a Substance Abuse Professional for a Substance Abuse Assessment (refer to the Substance Abuse Assessment Standard). Failing to meet with the Substance Abuse Professional or attend a scheduled Substance Abuse Assessment is a violation of this Policy.	Added	In fact, it is modified and not added; Corresponds to “After any confirmed violation, the employee may be referred for a SAE assessment to determine whether there is a need for a structured treatment program.” This is not detected by the toolset since the wording is significantly different

An alcohol test of .02 BAC, or higher, for anyone working at a designated dry site is considered a violation of this Policy. In all other situations, an employee who has an alcohol test result of .02 to .039 BAC in a reasonable cause or post incident situation will be removed from duty until considered safe to return (at a minimum not before their next work day or shift) and may be subject to corrective action.	Removed	Such specifics are removed from the corporate documents; the reason for removal may be to generalize but must be analyzed
Conditions for Continued Employment: Should the Company determine that employment will be continued after a violation of the Policy or Supporting Standards, the Employee will be required to enter into an agreement governing their continued employment which may require any or all of the following actions, or any other condition appropriate to the situation:	Modified	Implies that it is a necessary / general statement; "after a violation of the Policy or Supporting Standards" corresponds to "in a specific circumstance following a policy violation,"; paraphrasing implies that practice guideline has been properly referenced

There are two reasons for identical or modified portions. Either it is a general statement or a necessary statement. The modified portion represents the proper reference of the practice guideline. Although some are well-detected, others are not detected in Toolset 9 (Copyleaks). There are some errors from the software and others from complex sentence structure changes. Differences based on context or for specific domain are hard to detect for the machine. Toolset 10 (Copyscape) does not distinguish minor changes nor related meanings. However, it detects exact same three or more consecutive words.

Table 5-10: Document Conformance Map Created by Toolset 10 (Copyscape)

Practice Guideline	Matching Words	Corporate Document
4.0 SOCIAL SITUATIONS 4.1 Company Sponsored Social Events		8. SOCIAL SITUATIONS
In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community.	26 words	In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community.
		Subject to any site specific limitations,
Responsible alcohol use	3 words	responsible Alcohol use
is		may be
permitted at Company sponsored social functions	6 words	permitted at Company sponsored social functions

held away from Company premises, which must have the prior approval of a Vice President level or above, and will be conducted in accordance with the Company's hosting guidelines. Anyone who attends and consumes alcohol must not be returning to or going to work after the event. 4.2 Business Hosting		with appropriate prior approval. Alternative transportation arrangements will be made available when possible.
Consistent with the above, if alcohol is made available to	10 words	Consistent with the above, if Alcohol is made available to
		Company
guests in the course of conducting	6 words	guests in the course of conducting
		Company
business (e.g., restaurant	4 words	Business (e.g.,, restaurant
meeting),		meetings),
employees are expected to use	5 words	Employees are expected to use
		reasonable
judgment and be responsible in hosting others.	7 words	judgment and be responsible in hosting others,
5.0		and remain in compliance with the Policy and Supporting Standards. Procedures for hosting events are set out in the Social and Business Hosting Standard. 9.
CONSEQUENCES OF A POLICY VIOLATION	5 words	CONSEQUENCES OF A POLICY VIOLATION
5.1		a)
General Requirements Any violation	4 words	General Requirements: Any violation
of the provisions		
of this Policy	3 words	of this Policy
		and Supporting Standards
may result in	3 words	may result in
corrective action or		discipline up to and including
termination of employment.	3 words	termination of employment.
Management has the authority and discretion to hold out of service any individual who is believed to be involved in an incident that could lead to corrective action pending the results of the investigation. The appropriate action		In all situations, an investigation will be conducted to verify that a Policy or Standard violation has occurred. The appropriate discipline
in a particular case depends on the nature of the policy	11 words	in a particular case depends on the nature of the Policy
		or Standard
violation and the circumstances surrounding the	6 words	violation and the circumstances surrounding the
situation;		situation.
the severity of the violation	5 words	The severity of the violation
may		will

warrant entering the	3 words	warrant entering the
corrective action		discipline
process at different levels	4 words	process at different levels.
or termination of employment. For all employees any confirmed situation of drug trafficking on Company premises will result in termination of employment. A positive drug test, or an alcohol test result of .04 BAC or higher, or a refusal to test are all considered a violation of this Policy. An alcohol test of .02 BAC, or higher, for anyone working at a designated dry site is considered a violation of this Policy. In all other situations, an employee who has an alcohol test result of .02 to .039 BAC in a reasonable cause or post incident situation will be removed from duty until considered safe to return (at a minimum not before their next work day or shift) and may be subject to corrective action. After any confirmed violation, the employee may be referred for a SAE assessment to determine whether there is a need for a structured treatment program. 5.2		General violations of this Policy include: i. failure to comply with the Policy and Supporting Standards; ii. a positive Alcohol or Drug test (refer to the Alcohol and Drug Testing Standard); or iii. a Failure to Test. b) Referral for Assessment: After any confirmed positive Alcohol and Drug test, an Employee may be referred by Health and Wellness to a Substance Abuse Professional for a Substance Abuse Assessment (refer to the Substance Abuse Assessment Standard). Failing to meet with the Substance Abuse Professional or attend a scheduled Substance Abuse Assessment is a violation of this Policy. c)
Conditions for Continued Employment Should the Company determine that employment will be continued	13 words	Conditions for Continued Employment: Should the Company determine that employment will be continued
in a specific circumstance following a policy violation,		after a violation of the Policy or Supporting Standards,
the employee will be required to enter into an agreement governing their continued employment which may require any or all of the following actions, or any other condition appropriate to the situation: •	32 words	the Employee will be required to enter into an agreement governing their continued employment which may require any or all of the following actions, or any other condition appropriate to the situation:
		i. temporary removal from their position; ii.
adherence to any recommended treatment and aftercare program; •	8 words	adherence to any recommended treatment and aftercare program;
maintenance of sobriety and satisfactory performance on return to duty; •		iii.
successful completion of a	4 words	successful completion of a

fit for duty test; • ongoing unannounced testing for a period determined on a case by case basis taking into account the recommendations of the SAE; •		return to work Alcohol and Drug test; iv. ongoing unannounced follow-up Alcohol and Drug testing for the duration of their agreement; v.
adherence to any	3 words	adherence to any
		ongoing
rehabilitation conditions or requirements; and •	5 words	rehabilitation conditions or requirements; and
no further violations of the Policy		vi. no further Policy or Standard violations
during the monitoring period. Failure to meet the requirements of the agreement	12 words	during the monitoring period. Failure to meet the requirements of the agreement
during the monitoring period		
will be grounds for	4 words	will be grounds for
corrective action		discipline
up to and including termination	5 words	up to and including termination.
of employment as set out in the agreement.		

The conformance map of Toolset 10 (Copyscape) is more straightforward and easier to follow than Toolset 9 (Copyleaks). For example, “corrective action” and “discipline” are the terms used in the two documents. Although they both fit in the contexts, their meanings are slightly different. These words are detected clearly from the toolset. They are detected more easily in Toolset 10 (Copyscape). Slight differences are also detected more easily. Because of the differences between period (.) and semicolon (;), “situation.” and “situation;” are recognized as a difference. This can be useful in documents, such as contract or disclosure when meanings can significantly differ between punctuation marks. The difference between “will” and “may” are detected as well. They have a different level of enforcement. This piece of information may be important for auditors to have a sense of the force of the law of the policy.

5.2 Semi-Automated Benchmark Workflow Conformance Measurement Method and Toolsets

Along with the document conformance measurement method, the workflow conformance measurement method appears to have the potential to be semi-automated. Process mining software packages that allow process discovery and conformance checking functions may be candidates to achieve semi-automation in measuring the benchmark workflow.

There are not as many commercial process mining software packages compared to text mining software packages for document conformance measurement since process mining is a relatively new field. Thus, two software packages that provide academic versions are tested. Prices for the service are upon request. Both provide process discovery and conformance checking functions. Accuracy for process discovery and conformance measurement is checked, respectively.

5.2.1 Analysis of Event Log Preprocessing

The event log preprocessing module is necessary if the software returns errors during the data load. Common errors that must be addressed are related to timestamp formats. Typically, Microsoft Excel spreadsheets are accepted in the process mining software packages either in .xls or .csv format, or both. Excel has a "Format Cells" function and the user must know how to format the timestamps to avoid errors. Incorrect format designations may lead to wrong analysis

For example, for the following error "Error loading value 29/03/2020 18:13 (Row: 1, Col:2): Could not map data type. Object seen: 29/03/2020 18:13, Data Type: class java.lang.String, expected: DATE," date format must correctly be defined. If it is typed "MM/dd/yyyy HH:mm" instead of "dd/MM/yyyy HH:mm," the software packages do not run and return the error message or provide wrong analysis.

Another example error “Error loading value NULL (Row: 1, Col:7): Could not map data type. Object seen: NULL, Data Type: class java.lang.String, expected: DATE” is due to the null timestamps for completed time in an event log. Some software packages ignore the null timestamps, but others do not. Toolset 13 (LANA Labs) gives three options for this type of error. (1) ignore the value, (2) ignore the activity that includes this value, or (3) ignore the case that includes this value. However, Toolset 14 (Celonis) requires resolving the issue before entering the event log. Considering the fact that completed timestamp is not required to run the functions and also considering that for some software packages, the completed times do not affect the results, it is an issue that companies must fix. They can resolve this issue by either reflecting the completed timestamps precisely to the results or providing the results that are relatively inaccurate but still consistent and meaningful with entered data.

5.2.2 Analysis of Performance Measurement and Process Discovery Toolset Verification

Performance measurement is an optional module to gain additional insights from the conformance measurement analysis. A benchmark workflow shares a similar characteristic as a non-compulsory practice guideline. Thus, the benchmark workflow may be modified based on performance and conformance. Conformance results may provide enough information when the benchmark workflow is optimal and fixed. Thus, an auditor must consider whether to apply this module or not.

Depending on how the audited companies define key performance indicators (KPIs,) performance can be measured either manually or in automated ways. For instance, meeting the deadline for some activities may be a KPI. This may be done automatically or manually. “IF” function from Excel can be used to determine if the completed time is before the deadline in an automated way. (e.g., =IF(completed time<deadline, “S”, “F”); which means if completed time is less than deadline print “S”, otherwise print “F”) The user must remember that in Excel, every date and time combination has

“general” format (e.g., 3/10/2011 5:12:00 PM = 40612.7166666667). Thus, it does make sense to have such IF formula.

Once performances are measured and recorded, with the “Filter” function, they can be easily grouped as desired. For example, if “S” or “F” is assigned each workflow instance, they can be divided into “S” group and “F” group. Once they are separated in the new spreadsheets, they are ready to be processed with the software packages.

Process discovery function is an optional function when workflow instances are independent and not grouped, and manual conformance measurement takes place with an event log. The process discovery function has an ability to generalize the event log data and to visually illustrate in a flowchart form. It can also group workflow instances and integrate the information into one flowchart. However, in the process of it, information loss or misinterpretation may occur. For example, when the created timestamp and the completed timestamp of activity are the same, and if there are multiple activities with the same situation, a human knows that these activities are created at the same time and completed at the same time instantly. However, for the machine to compute this situation, it may be tricky, and the machine may return the results showing the activity has happened repeatedly or sequentially, when it actually happened simultaneously and instantly. Though it is not an incorrect analysis, it does not explain the reality that can be presented easily in manual ways.

Process discovery toolsets are measured in terms of accuracy. Accuracy is defined as how a workflow instance is presented correctly and how multiple workflow instances are integrated correctly. This is verified with an example event log data. The criteria for accuracy are summarized in Table 5-11.

Table 5-11: Legend for Accuracy Criteria for Process Discovery Tools

Rating of Accuracy	Definition
A	Successfully returns the visual representations, mostly accurate (e.g., considering completed timestamps) for both individual workflow instances and workflow instance groups
B	Successfully returns the visual representations, mostly accurate (e.g., considering completed timestamps) for only individual workflow instances and not workflow instance groups
C	Successfully returns the visual representations, but mostly not accurate (e.g., not considering completed timestamps)
D	Unsuccessful in returning the visual representations

The event log data has two workflow instances with 53 activities. The two workflow instances are grouped as one group for this experiment. How the software packages visualize each workflow instance and integrate two workflow instances as one flowchart are the criteria for the verification. The event log data used for the verification is attached in Appendix E. The results for an individual workflow instance are illustrated in Figure 5-1. Workflow instance 27 is used to derive the result.

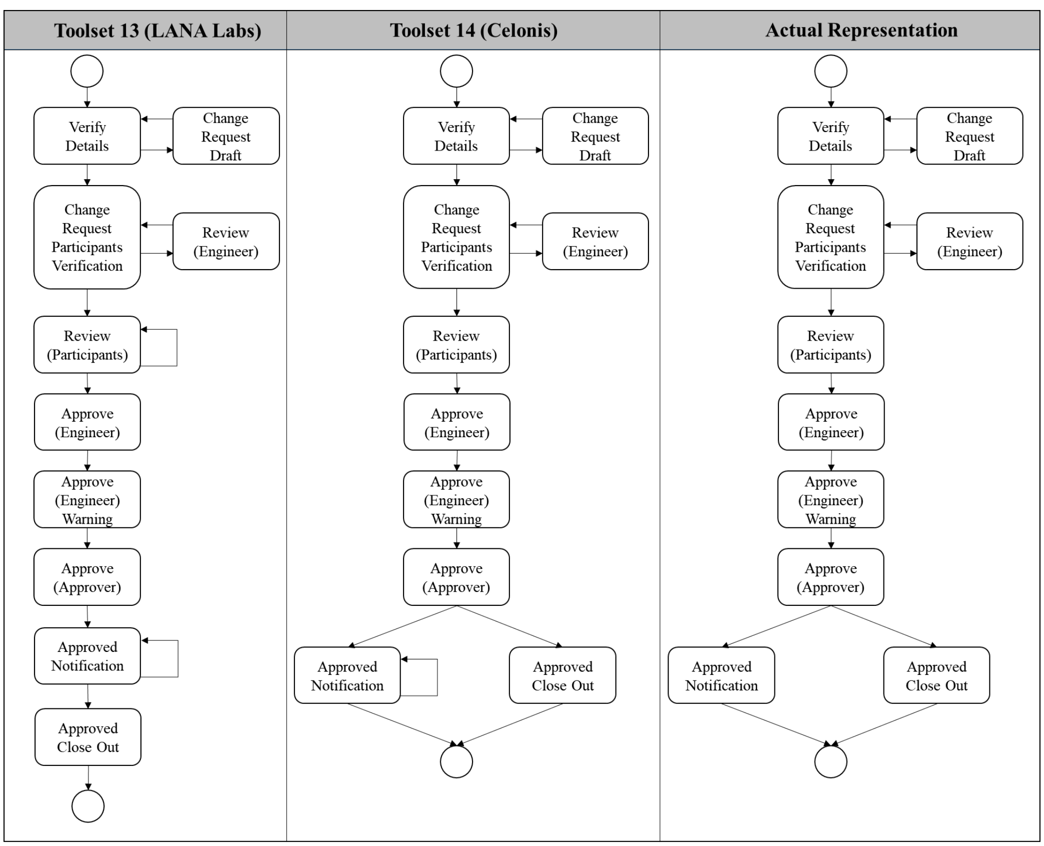


Figure 5-1: Accuracy Comparison among Software Packages for an Individual Workflow Instance

Toolset 14 (Celonis) is closer to the actual representation. It considered completed timestamps of activities such as "Review (Participants)." The completed timestamps of multiple "Review (Participants)" differ. However, they all started the same and ended before "Approve (Engineer)." Thus, "Review (Participants)" must not be a repetitive activity. Toolset 14 (Celonis) addresses this issue correctly. Moreover, the created timestamps of "Approved Notification" and "Approved Close Out" are the same. It could be interpreted as sequential since the completed timestamp of "Approved Notification" is the same as the created timestamp. However, it is not preferred since these two activities started at the same time after "Approve (Approver)." Thus, Toolset 14 (Celonis) represents this issue more closely to actual.

A workflow instance group is also tested. Workflow instance group is the combination of Workflow instance 26 and Workflow instance 27. Changes are marked in red arrows in Figure 5-2. The change is from adding Workflow instance 26 to the results from Workflow instance 27.

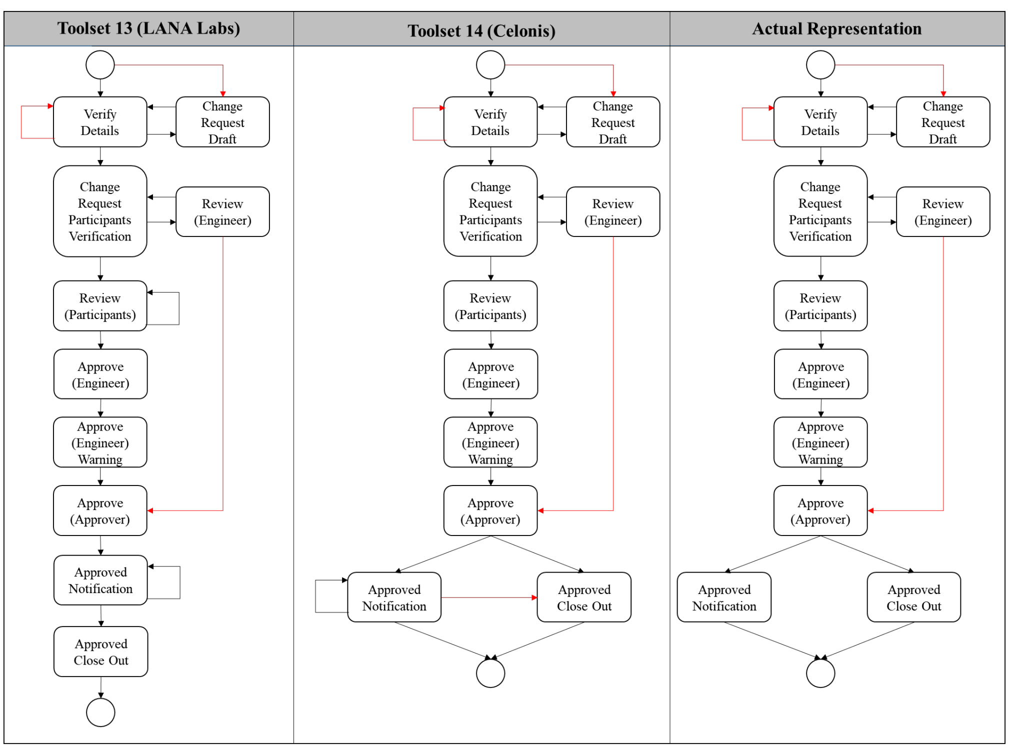


Figure 5-2: Accuracy Comparison Among Software Packages for a Workflow Instance Group

Toolset 13 (LANA Labs) incorporated the same changes as actual representation. Toolset 14 (Celonis) added a transition from “Approved Notification” to “Approved Close Out” which is unnecessary.

The summary of accuracy for the process discovery function is presented in Table 5-12. Cost refers to the price of the software packages. Labor costs are not included. Duration refers to computation time. Accuracy and strength of each toolset is revealed.

Table 5-12: Verification Results for Process Discovery Toolsets

Criteria	Toolset 13 LANA Labs	Toolset 14 Celonis
Cost	Price upon request	Price upon request
Duration	1 sec	1 sec
Accuracy	B	B
Strength	Workflow instance groups representation	Individual workflow instances representation

The price of the process mining software packages depends on scale of a company and functions that the company desires to use. Whether there is one workflow instance or 10,000 workflow instances, the computation time is quick enough to conclude that there is not much difference. Although the results are not 100% as intended, the toolsets returned results with adequate accuracy.

The resulting flowcharts of workflow instance groups are called discovered workflows in this study. These discovered workflows are used for the conformance checking function in the conformance measurement module.

5.2.3 Analysis of Conformance Measurement and Conformance Checking Toolset Verification

The conformance measurement module (Module 3) is an essential part of the capital project practice benchmark workflow conformance measurement method. This module can be semi-automated with software packages. Generally, software packages convert an event log into a discovered workflow and then compare the discovered workflow against the benchmark workflow. This seems accurate, but, when the discovered workflow derived from the performance measurement module (Module 2) is incorrect, the conformance measurement result also is inaccurate. Moreover, the software packages typically do not allow the users to define conformance but have their own metrics. Thus, manual conformance measurement may be the only option. Manual conformance measurement also allows crosschecking with the event log. Thus, the accuracy of the manual process may be higher than the automated ones. However, the challenge of manual computation is the manhours it takes to analyze all the workflow instances. It may not take long to analyze a few workflow instances. However, there are hundreds and thousands of workflow instances in an event log. Comparing durations between humans and machines is meaningless since software packages take less than a minute to process 500 workflow instances while it takes a minute or more for humans to check one workflow instance.

To verify the accuracy of the conformance checking function from software packages, the ratings of accuracy are defined in Table 5-13. The same event log that has been used for process discovery software packages verification is used for conformance checking software packages verification. The event log is provided in Appendix E.

Table 5-13: Legend for Accuracy Criteria for Conformance Checking Tools

Rating of Accuracy	Definition
A	Successfully computes the comparison of the discovered workflow against the benchmark workflow; Correctly detects all the non-conformances with no errors
B	Successfully computes the comparison of the discovered workflow against the benchmark workflow; Correctly detects most of the non-conformances with minor errors
C	Unsuccessfully computes the comparison of the discovered workflow against the benchmark workflow; Either detects false non-conformances or fail to detect non-conformances
D	Fails to compute the comparison of the discovered workflow against the benchmark workflow

An appropriate benchmark workflow is necessary for the conformance checking. The actual representation is used as the benchmark workflow for the comparison. Figure 5-3 is the Business Process Model and Notation (BPMN) format representation.

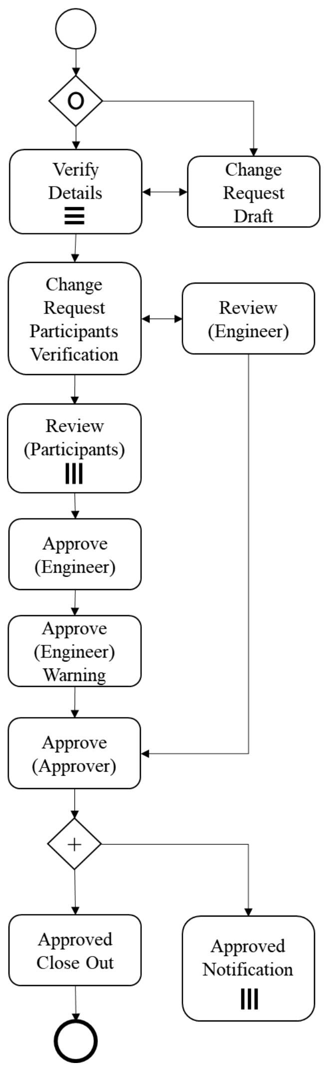


Figure 5-3: The Benchmark Workflow (BPMN Format)

The BPMN has its own gateway symbols other than activity symbols. There are event types (start event symbol (◯): signals the first step of a process, end event symbol (**O**): signals the final step of a process), activity symbol (task symbol (): the most basic level of an activity), and gateway symbols (exclusive symbol (): breaks the flow into one or mutually exclusive paths based on condition, parallel symbol (): concurrent tasks). Alternatively, inclusive symbol () can be used to break the flow into one or more flows. The reason for the use of exclusive symbol rather than the inclusive symbol was due to the sequential activities that occurred in the event log. The parallel sign (|||) indicates that multiple participants may execute the activity concurrently. The sequential sign (≡) indicates that the activity may be completed consecutively. The comparison between the conformance checking software packages is summarized in Table 5-14.

Table 5-14: Verification Results for Conformance Checking Toolsets

Criteria	Toolset 13 LANA Labs	Toolset 14 Celonis
Cost	Price upon Request	Price upon Request
Duration	1 sec	1 sec
Accuracy	D	C

Toolset 13 (LANA Labs) was not able to process the benchmark workflow input. Toolset 14 (Celonis) was able to process the benchmark workflow; however, the output is inaccurate. The output of Toolset 14 (Celonis) is summarized in Table 5-15. These non-conformance types are false according to the event log. Based on the benchmark workflow and the event log data, all the workflow instances must be returning "conformance." However, Toolset 14 (Celonis) detected four non-conformance types assigning the workflow instances as "non-conformance."

Table 5-15: Non-Conformance Types Detected by Toolset 14 (Celonis)

No.	Non-Conformance Types	Number of Workflow Instances (WF_ID)
1	"Verify Details" is followed by "Change Request Participants Verification"	2 (26, 27)
2	"Review (Engineer)" is followed by "Approve (Approver)"	1 (26)
3	"Change Request Draft" executed as START activity	1 (26)
4	"Verify Details" executed as START activity	1 (27)

These false detections are due to software bugs or the complexity of workflow. Especially, Toolset 14 (Celonis) does not have strength when workflow instance groups are formed. Conformance checking software packages do not have adequate accuracy. Thus, the toolsets are not further used for conformance measurement in Chapter 6.

5.2.4 Analysis of Workflow Conformance Map

Once the performance measurement and conformance measurement modules are completed, it can be decided whether to keep or modify the original benchmark workflow. The decision is based on performance and conformance results. If the performance of a case group is successful when the discovered workflow of the case group is in conformance, the auditor may suggest keeping the original benchmark workflow. If the performance is unsuccessful when the discovered workflow of the case group is in conformance, the auditor may suggest modifying the original benchmark workflow. If the performance is successful when the discovered workflow of the case group is in non-conformance, the auditor may suggest modifying the original benchmark workflow by cautiously accepting the non-conformance types as conformance. If the performance is unsuccessful when the discovered workflow of the case group is in non-conformance, the auditor may suggest enforcing the original benchmark workflow but not make any conclusions. In the simplified example from the conformance measurement module (Section 5.2.3), the conforming discovered workflow was estimated as non-conformance by Toolset 14 (Celonis). This result reverses the conclusions that the auditor may make. Thus, this is considered a major toolset error.

Since there is no function in the software packages to create a workflow conformance map as defined by this study, this module is considered as the module that can be completed manually. The workflow conformance map is a visualization of the discovered workflow over a benchmark workflow. The map visually emphasizes the conformance between actual and planned workflows and guides the auditor.

Workflow conformance maps can be created by illustrating the conformance measurement results. Workflow conformance maps may be utilized to make decisions on whether to keep or modify the original benchmark workflow with performance measurement results.

The example of a conformance map is illustrated in Figure 5-4. Workflow instance 26 is illustrated on the benchmark workflow from the previous subchapter. Activities in conformance are marked in gray background and blue boundary. Flow paths in conformance are marked in blue arrows. Non-conformances are supposed to be marked in red; but, in this case, every activity and flow path are in conformance. Note that when undefined activities or flow paths are added, this must be marked.

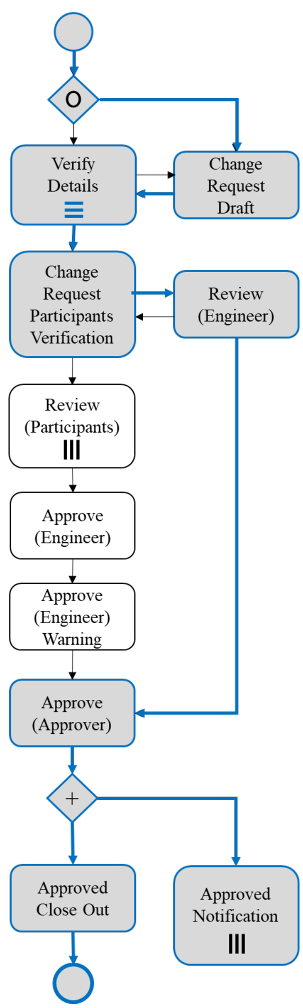


Figure 5-4: An Example of Workflow Conformance Map (Workflow Instance 26)

In some cases where modifications of the benchmark workflow are challenging, one way to modify the benchmark workflow is by finding the discovered workflows that have resulted in good performances. Even the discovered workflows that are not in conformance or the paths that have happened less frequently may be candidates.

5.3 Discussion and Conclusions of the Framework Analysis and Toolset Verification

To achieve semi-automated capital project practice conformance measurement, document and workflow components were extracted from the practice conformance model from Chapter 3. These two components have their own methods and they were defined in Chapter 4. Each subchapter discussed why the corresponding modules are included and how they can be semi-automated. Whether they are optional or not was addressed as well as the possible orders of modules. For example, in the document conformance measurement method, module 2, module 3, and module 4 were optional. Although at least one module must be selected to identify the method as a practice conformance measurement method, the modules and the order may differ depending on how an auditor defines the practice conformance. The results of module 2, module 3, and module 4 of document conformance measurement are likely to have formats with different levels of details. Considering the accuracy of the toolset verification results, Toolset 4 (WPS) and Toolset 9 (Copyleaks) are selected for Chapter 6 case studies of semi-automated document conformance measurement. Considering the accuracy and capabilities of the toolsets, Toolset 14 (Celonis) is selected for the case study of semi-automated workflow conformance measurement.

With framework analysis, toolsets for some modules are introduced and evaluated in terms of cost, duration, and accuracy. Table 5-16 and Table 5-17 summarize capabilities, limitations, and potential uses of the toolsets.

Table 5-16: Document Conformance Measurement Toolset Analysis

Toolset	Module	Capabilities	Limitations	Potential Uses
Toolset 1 (Google OCR)	Text Extraction	• Converts .jpeg, .png, .gif, or PDF file formats • Bold, italics, font size, font type, and line breaks are likely to be retained • Cursive handwritings are partially detected	• Lists, tables, columns, footnotes, and endnotes are likely not to be detected • Images are likely not to be retained	• Handwritten document conversion to machine-readable documents
Toolset 2 (Microsoft Word)	Text Extraction	• Converts PDF file documents into editable Word format • Bold, italics, font size, font type, and line breaks are likely to be retained	• Lists, tables, columns, footnotes, and endnotes are likely not to be detected • Handwritten document conversion not supported (OCR feature in OneNote)	• Simple PDF conversion to editable file formats (without any images or columns)
Toolset 3 (Adobe Pro)	Text Extraction	• Converts PDF file documents into editable Word format • Bold, italics, font size, font type, and line breaks are likely to be retained • Lists, tables, columns, footnotes, and endnotes are likely to be retained • Images are likely to be retained • The navigation panel is likely to be retained	• Handwritten document conversion is supported, but the accuracy is low	• Complex PDF conversion to editable file formats (with images and tables)
Toolset 4 (WPS)	Text Extraction	• Converts PDF file documents into editable Word format • Bold, italics, font size, font type, and line breaks are likely to be retained • Lists, tables, columns, footnotes, and endnotes are likely to be retained • Images are likely to be retained • The navigation panel is likely to be retained	• Handwritten document conversion is supported, but the accuracy is low	• PDF conversion to editable file formats • (with images and tables)
Toolset 5 (WordClouds)	Keyword Extraction	• Accepts: paste/type text; text file; URL; MS Office document; PDF document	• Unable to select to ignore stop words or numbers	• PDF files or MS Office documents

		• Possible to save an image, and edit result • Possible to select to ignore stop words • Sorts in terms of word frequencies or alphabetical order	• Stemming and lemmatization unavailable • English only	
Toolset 6 (WordArt)	Keyword Extraction	• Accepts: paste/type text; .csv, .xls; URL • Possible to select to ignore stop words or numbers • Allows stemming • Possible to save an image, and edit result in .png, .jpeg, .svg, .pdf , .html with cost • Possible to search words	• Sorting unavailable • English only	• Keyword extraction with specific stop words • Keyword extraction that requires stemming
Toolset 7 (TagCrowd)	Keyword Extraction	• Accepts: paste/type text; text file; URL • Language options • Possible to select the maximum number of words to show or minimum frequency • Possible to select to show frequencies • Stemming and lemmatization possible • Possible to select case sensitivity • Possible to remove unwanted words • Possible to save image as .html or PDF	• Unable to show stop words • Sorting unavailable	• Different languages (French, Spanish, etc.) • Keyword extraction that requires stemming and lemmatize-ation
Toolset 8 (WordItOut)	Keyword Extraction	• Accepts: paste/type text; tables in spreadsheet • Possible to view and edit default stop words (filtered words, punctuation characters) • Possible to save image as .png	• Sorting unavailable • Stemming and lemmatization not available • English only	• Keyword extraction that controls stop words
Toolset 9 (Copyleaks)	Text Matching	• Accepts: paste/type text; files (e.g., PDF; Microsoft Office); URL • Possible to compare multiple documents at once • Possible to select to identify identical, minor	• Does not count every portion of words that match • Returns only one matching rate per use (Require selecting one document as	• PDF, Word format file comparison • Compares multiple documents at once

		changes, and related meaning words • Possible to download the result in PDF • Possible to detect regardless of the order of texts • Provide overall matching rate and similar words	original and others suspect)	• Identifies synonyms or minor changes
Toolset 10 (Copyscape)	Text Matching	• Accepts: paste/type text • Highlights and counts every portion of words that match • Counts the total number of words for both texts and provides two numeric values as the matching rate at once; also, the total number of matching words	• Unable to detect matching if not in order • Not possible to compare multiple documents (only two) • Finds every minor change as not matching; unable to separate minor changes (e.g., punctuations, numbering) • Not possible to download the results	• Simple text comparison • Detecting exact matches is enough (no need for identifying minor changes)
Toolset 11 (PlagScan)	Text Matching	• Accepts: paste/type text; files (e.g., PDF; Microsoft Office); URL • Possible to compare multiple documents at once (including all other sources that may be related) • Possible to download results in PDF or Word document with annotations	• Unable to compare only specific documents (web-based)	• Other unknown sources may have been referenced
Toolset 12 (Countwords free)	Text Matching	• Accepts: paste/type text • Comparison results can be merged in one by highlighting removed and added portions • Possible to download the results in PDF or Word document	• Unable to detect matching if not in order • Not possible to compare multiple documents (only two) • Finds every minor change as not matching; unable to separate minor changes (e.g., punctuations, numbering)	• Simple text comparison • Detecting exact matches is enough (no need for identifying minor changes)

			• Not possible to download the results	

There are strengths of applying existing commercial software packages for the semi-automated document conformance measurement method. Although they are created for their own purposes, applying them ensures some level of accuracy especially when they are paid services. There are exceptions such as Toolset 1 (Google OCR) for handwriting recognition. Toolset 1 (Google OCR) is open-source; however, it has better accuracy than other software packages.

Commercial software packages are also founded upon theories, principles, and experiments. They are used for industrial applications. Moreover, the software packages have improved in terms of accuracy and functionalities over the years. There still exist technical limitations, but these document conformance measurement software packages are improving quickly.

Table 5-17: Workflow Conformance Measurement Toolset Analysis

Toolset	Module	Capabilities	Limitations	Potential Uses
Toolset 13 (LANA Labs)	Performance & Conformance Measurement	• Process discovery function: filters to simplify or specify the discovered workflow • Conformance checking function: Possible to view every activity status (conform, skipped, inserted) within every case • Possible to create a benchmark workflow; Possible to filter by case • Possible to process "null" timestamps; options must be selected	• Technical limitation such as unclear timestamp format • Unable to edit benchmark workflows • Not every benchmark workflow can be uploaded • Transition non-conformances are difficult to identify • Cases that completed timestamps are incorrectly addressed	• Simple and straight forward data to create a discovered workflow or check general conformance against default benchmark workflow or changes
Toolset 14 (Celonis)	Performance & Conformance Measurement	• Process discovery function: structured filters to simplify or	• Unable to process "NULL" timestamps; event	• Complex large data to create a discovered workflow or

		specify the discovered workflow • Conformance checking function: Possible to edit the benchmark workflow • Possible to change non-conformance type into conformance manually by enlisting in the whitelist • Possible to create a benchmark workflow • Possible to filter by case • Transition non-conformances are clear	log must be preprocessed • Cases that completed timestamps are incorrectly addressed	check conformance • Freely edit the benchmark workflow • Edit conformance types manually

When an audited company does not have a specific benchmark workflow (or a formal process) to start with, potentially, the company can execute the process and collect initial event log data to create a benchmark workflow model using process discovery function. Filters from the process discovery function can be used to create a benchmark workflow or improve a rough one. The filters are originally intended for an auditor to visualize the frequency of activities and transitions. While the discovered workflow remains connected, some activities and transitions may not appear when filters are adjusted. It simplifies or specifies the benchmark workflow.

Semi-automated benchmark workflow conformance measurement toolsets are reasonable when there is a massive amount of data. An auditor must know the process to resolve some incorrectly identified non-conformances. Compared to semi-automated document conformance measurement toolsets, the workflow conformance measurement toolsets are likely to be more costly taking longer learning curve with lower accuracy due to the complexity of both data and software packages. Therefore, it is suitable for large capital projects with complex workflows and massive data. Otherwise, the decision to adopt the software package may be unsuitable.

Chapter 6 Validation of the Semi-Automated Capital Project Practice Conformance Measurement Framework

From the practice conformance model (Chapter 3), documents and workflows were perceived to have the potentials to be measured in semi-automated ways. In this chapter, the framework from Chapter 4 and the semi-automated toolsets from Chapter 5 are implemented with the real data attained from companies in the construction domain. Experiments are designed and analyzed for the validation of the framework.

Seven methods (DM, DS1, DS2, DS3, WM, WS1, and WS2) were developed and tested to compare cost, duration, and accuracy (error rate) between manual and semi-automated methods. Automation levels were adjusted to compare the effectiveness of the module combinations (Figure 6-1).

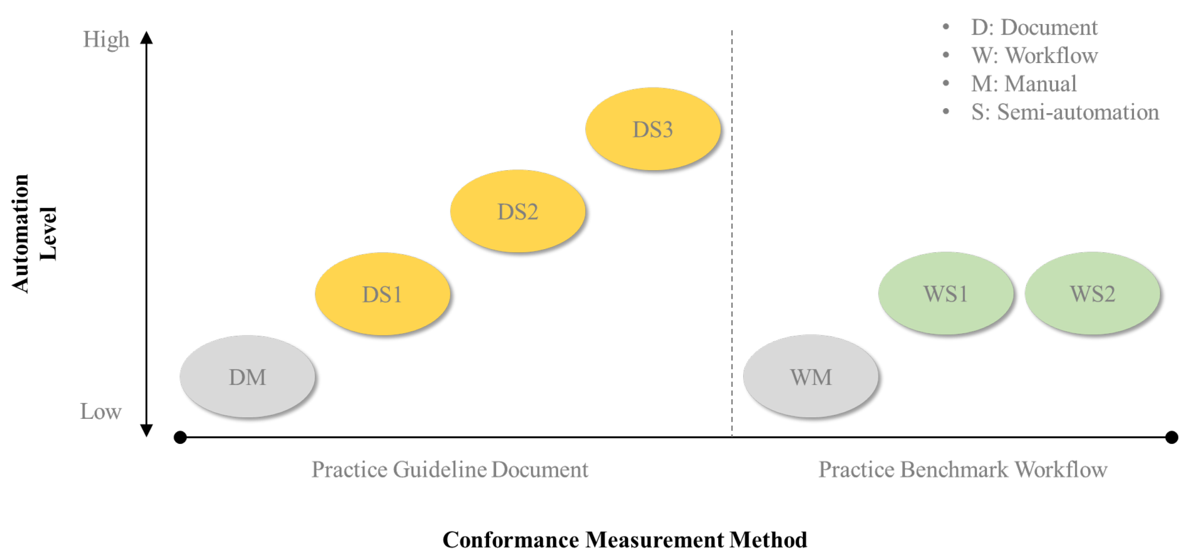


Figure 6-1: Automation Levels of Practice Conformance Measurement Methods

6.1 Validation of the Practice Guideline Document Conformance Measurement Method

Four methods (DM, DS1, DS2, DS3) were applied to five representative cases. For each case, the outcomes of the four methods were different. This may be due to different factors such as the number of pages of documents, the topics of the documents, and the relationships between two documents. However, in general, the semi-automated methods took less time while increasing conformance measurement accuracy. As the number of pages increased, duration gaps between manual method (DM) and the semi-automated methods (DS1, DS2, DS3) became more apparent.

6.1.1 Implementation of the Practice Guideline Document Conformance Measurement Method

For the validation, the practice guideline document conformance measurement flowchart was applied (Figure 4-3). Additionally, functional requirements were followed (Figure 4-2). One of the differences between the flowchart and the functional requirements is that the flowchart includes the context to what auditors must follow before and after the semi-automated document conformance measurement (e.g., prepare documents, conduct an audit analysis). The flowchart also provides a potential for customization, depending on the definition of conformance, by allowing different combinations of modules. On the other hand, the functional requirements specify the steps with examples and address potential automation more clearly. Therefore, the flowchart and functional requirements cooperatively functioned for the validation of the document conformance measurement method, as a part of the semi-automated practice conformance measurement framework.

According to the flowchart (Figure 4-3), first, documents were selected. Among the twenty-two documents that were attained, two documents were selected for the experiment. Then, the term conformance was defined. Based on the definition, the modules and the sequences of the modules can be determined. For example, if conformance was defined as the existence of specific conformance elements from a practice guideline, the information type detection module (Module 2) may be included. If conformance was defined as the inclusion of keywords, the keyword extraction module (Module 3) may be included. If conformance was defined as the inclusion of matching text portions, the text matching module (Module 4) may be included.

Figure 6-2 visually summarizes the possible definitions of conformance between two text portions. For this document conformance measurement validation experiment, matching, the identical portions of texts, was considered as conformance (❷). It was assumed that the same language (e.g. acronyms) was used across the documents (❸). Thus, all methods (DM, DS1, DS2, DS3) included the text matching module (Module 4).

Figure 6-2: Term Definitions and Relationships: "Conformance," "Matching," and "Related"

Four relationships are identified in Figure 6-2. The first relationship (❶) is where two text portions are related, in conformance, but not matching. They are not matching due to punctuation differences, redactions, or usage of synonyms. However, the text portions conform to each other since they semantically align. The second relationship (❷) is where two text portions are identical in form and meaning. The third relationship (❸) is where two text portions are identical in form but not semantically. For instance, homonyms may look the same; however, they have distinctive meanings. Acronyms and abbreviations are examples that may be different when expanded in full names. The last relationship (❹) is where two text portions are related but neither matching nor in conformance. An example would be added negations. Two text portions may be related but, once negation is added, they are not identical nor in conformance. Instead, they may be indicating the complete opposite of the other text portion.

The evaluation measures for Information Retrieval (IR) are used to assess how well the search results satisfied the user's query intent. The metrics can be applied to program the text matching software to "match," "conform," or "relate" two text portions. Metrics measure relevance, that is, how likely each result (retrieved) is to meet the information needs (relevant) of the user. By applying the evaluation of a user interaction (e.g. click/dwelling time) within the search system (online metrics) or measurement of judges' scores of the relevance (offline metrics), conformance can be measured. For offline metrics, judges score binary or multi-level relevance. Once conformance was defined, four methods, each including modules and sequences, were tailored. Four methods are listed in Table 6-1.

Table 6-1: Method Descriptions for the Application Experiment

Method	Automation Level	Selected Toolset	Sequence of Modules	Details
Manual (DM)	Level 0	• N/A	Module 4 (Manual) → Module 5 (Manual)	• Manually complete Module 4 and Module 5
Semi-Automation (DS1)	Level 1	• Toolset 2 (Microsoft Word)	Module 3 (Manual) → Module 4 (Auto) → Module 5 (Manual)	• Manually complete Module 3 and Module 5 • Automatically complete Module 4
Semi-Automation (DS2)	Level 2	• Toolset 4 (WPS) • Toolset 9 (Copyleaks)	Module 1 (Semi) → Module 2 (Manual) → Module 4 (Auto) → Module 5 (Auto)	• Semi-automatically complete Module 1 • Manually complete Module 2 • Automatically complete Module 4 and Module 5
Semi-Automation (DS3)	Level 3	• Toolset 9 (Copyleaks)	Module 4 (Auto) → Module 5 (Auto)	• Automatically complete Module 4 and Module 5

*Module 1: Text Extraction; Module 2: Information Type Detection; Module 3: Keyword Extraction; Module 4: Text Matching; Module 5: Document Conformance Map

The application experiment involved four methods (DM, DS1, DS2, DS3). Four methods selectively included text extraction (Module 1), information type detection (Module 2), keyword extraction (Module 3), and/or text matching (Module 4) modules. Five sets of cases iterated these four methods.

An automation level is defined as a perceived degree of automation involved in each automation method. Level 0 refers to the level to complete the manual processes, whereas Level 1, Level 2, and Level 3 refer to the levels that involve some degree of automation. The higher the level is, the lesser the manual work is involved relatively. For this experiment, the semi-automated method 1 (Level 1) included some automation compared to the manual method (Level 0). The semi-automated method 2 (Level 2) involved more automation than the semi-automated method 1 (Level 1). The semi-automated method 3 (Level 3) was considered the highest automation level since fewer manual steps (e.g., information type detection module (Module 2)) were involved. Based on the methods, the toolsets were selected accordingly.

The Society of Automotive Engineers (SAE) also defined levels of automation (NHTSA, 2020). There are six automation levels from no automation (Level 0) to full automation (Level 5). In between, levels 1–4 exist as semi-automation: driver assistance (Level 1), partial automation (Level 2), conditional automation (Level 3), and high automation (Level 4). The categorization is used to describe which level the industry is at. Likewise, clearly defined automation levels can be used to describe how much automation a company has reached. The automation level can also be an indicator to update the current status of the technology and to pursue next step if desired.

The manual method (DM) was designed to achieve the same end result from the other methods; thus, it may be unlike the real processes that an auditor follows. The manual method (DM) skipped the text extraction (Module 1), information type detection (Module 2), and keyword extraction (Module 3) modules. This was not because these modules were less important but because they were innately included in the following module (the text matching module (Module 4)) when performed manually. In summary, the manual method (DM) was completed by manually comparing two documents and highlighting the matching portions.

The first semi-automated method (DS1) was designed to include some automation to the manual method (DM). The document conformance map module (Module 5) was completed manually the same as the manual method (DM). The keyword extraction module (Module 3) was added to the manual method (DM) but was completed manually, not using any keyword extraction software package. The reason for not using the keyword extraction toolsets was due to the low accuracy of the relevant toolsets. The text matching module (Module 4) was completed automatically by applying the keywords identified from the keyword extraction module (Module 3). In the text matching module (Module 4), the search function that is typically embedded in electronic document file viewers were utilized. In Word and PDF viewer, the search function can be found using the "Find and Select (shortcut key: Ctrl+F)" feature. In summary, the keywords identified from one document were searched in the other document and the matching results were highlighted.

The second semi-automated method (DS2) was designed to include the information type detection module (Module 2) and to automate the text matching module (Module 4) and the document conformance map module (Module 5). To include the information type detection module (Module 2), the text extraction module (Module 1) was a prerequisite. After the information type detection module (Module 2) was completed, the selected information type ("description/narratives" conformance element) was compared in the text matching module (Module 4). By using the software package (Toolset 9 (Copyleaks)), the text matching module (Module 4) and the document conformance map module (Module 5) were completed automatically.

The last semi-automated method (DS3) was designed so that the information type detection module (Module 2) and the keyword extraction module (Module 3) were skipped while the text matching module (Module 4) and the document conformance map module (Module 5) were completed automatically. The semi-automated method 3 (DS3) was expected to reduce the total duration since it only included two modules.

Once a method was selected, the modules were executed in sequence with some toolsets if needed. The end result (i.e., the document conformance map) was generated so that an auditor can begin analysis. A screenshot of application experiments of the document conformance measurement methods is presented in Figure 6-3. By using software (DS2 & DS3) or manually (DM & DS1) conforming portions were marked.

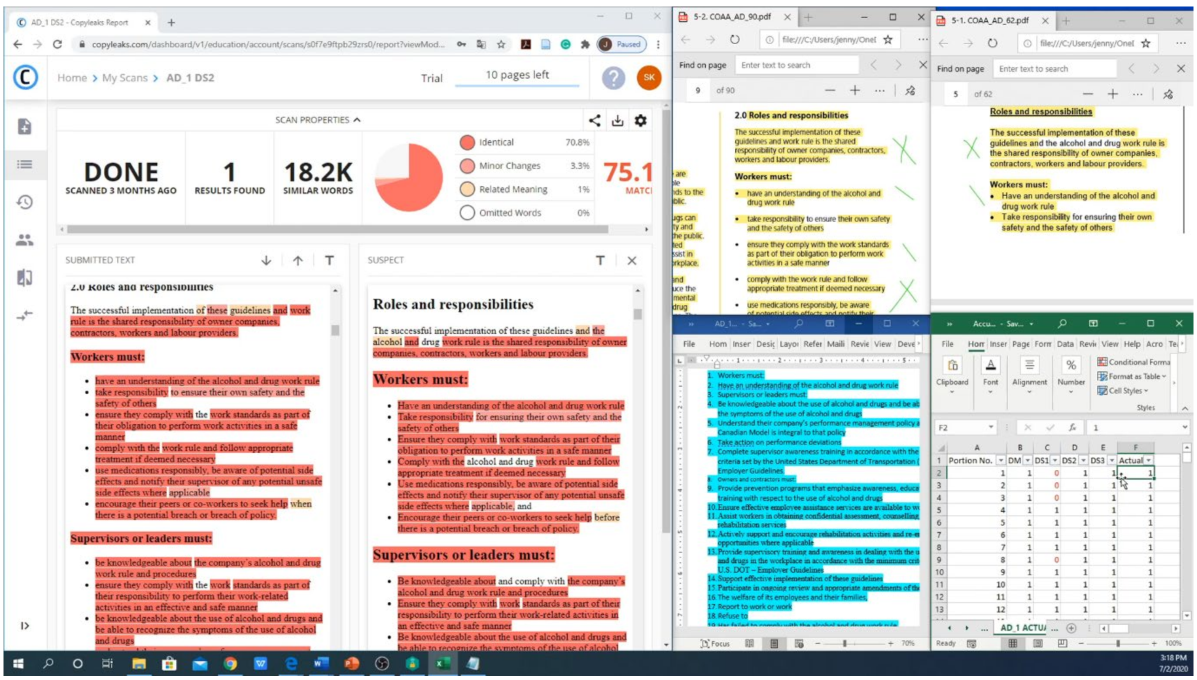


Figure 6-3: A Screenshot of Text Matching Software (Left: Red), Manual Text Matching (Top Right & Center: Yellow), Matching Portions Extracted (Bottom Center: Blue), Record (Bottom Right)

6.1.2 The Results and Analysis of the Practice Guideline Document Conformance Measurement Method

The main objectives of the application experiment are to validate the efficiency, cost-effectiveness, and correctness by comparing duration, cost, and accuracy of each method. To validate the document conformance method, over 120 cases were studied. For the case studies, three topics were selected and a total of twenty-two documents were obtained from seven companies, three institutions, and a government body. Among 120 cases that were crosschecked manually, five representative cases were selected for the experiment (Table 6-2). Manual crosschecking was necessary in order to select representative cases.

Selected conformance topics includes: (1) alcohol and drug-related, (2) project management, and (3) interface management documents. They are all critical documents for successful capital project management. Alcohol and drug-related documents were received from four companies, an institution, and a government body. Project management documents were received from three companies and an institution. Interface management documents were received from a company. For every case, four methods (DM, DS1, DS2, DS3) were applied. The application results were compared.

Table 6-2: The Number of Documents and Cases

Topic	The Number of Documents	The Number of Cases (DM: DS1: DS2: DS3)	The Number of Representative Cases
Alcohol and Drug (AD)	16	120 : 120 : 3 : 3	3
Project Management (PM)	4	6 : 6 : 1 : 1	1
Interface Management (IM)	2	1 : 1 : 1 : 1	1
Total	**22**	**127 : 127 : 5 : 5**	**5**

The reason for conducting 127 experiments manually (DM) and semi-automatically (DS1) was to find the representative cases. After finding representative cases, the other the semi-automated methods (DS2, DS3) were applied. The five representative cases are summarized in Table 6-3.

Table 6-3: The Descriptions of Representative Cases

Case ID	Topic	Source	Characteristic	Pages
AD_1	Alcohol and Drug	Same institution	Revision	90pg, 62pg
AD_2	Alcohol and Drug	Same company	Similar formats	5pg, 3pg
AD_3	Alcohol and Drug	Institution & Company	Company adapted institution	62pg, 5pg
PM_1	Project Management	Two companies	Two projects	94pg, 31pg
IM_1	Interface Management	Same company	Redacted	36pg, 36pg

Each representative case was unique because of its source, characteristic and pages of documents. The first case (AD_1) was comparing two versions from the same institution. Some changes appeared in these two versions. This was the case with the longest expected duration for the manual method (DM) because the conformance measurer had to pay attention to find the differences from similar texts. These two documents were also relatively longer than the others.

The second case (AD_2) was the comparison between two documents from the same company. Because the documents were from the same company, they had a similar introduction. The documents were relatively shorter. Though differences may not be difficult to identify, human errors were expected since the conformance measurer may overlook the shorter documents.

The third case (AD_3) was the comparison between an institution and a company. This case was a good example to demonstrate document conformance measurement between two sources. Especially, since the company adapted a part of the institution document, the matching result revealed the company's intentional adaption.

The documents for the fourth case (PM_1) were obtained from two companies. Unlike alcohol and drug-related documents, which were mostly guidelines or policies, the project management documents were plans for specific projects. The texts from a project document were unlikely to match with texts from other project management documents. Because of the definition of conformance, the result was expected to indicate that these two documents do not match at all. However, if conformance was defined differently (e.g., the existence of information types), the result would have been different.

The last case (IM_1) was the comparison between two interface management documents from the same company. One document had been redacted. However, failure to redact some names and titles of the redacted version clarified that not much effort was put in for the redaction process. This case was another good example of conformance measurement to detect whether two sources had been randomly copied. The final results are summarized in Table 6-4 and explained in detail in the subsequent sections.

Table 6-4: Summary of the Application Experiment

Case ID	Criteria	Manual (DM)	Semi-Automation 1 (DS1)	Semi-Automation 2 (DS2)	Semi-Automation 3 (DS3)
		Level 0	Level 1	Level 2	Level 3
AD_1	Duration (hr)	4.70	1.03	0.58	0.02
	Cost (USD)	$846.00	$186.00	$20.66	$8.21
	Accuracy (%)	77.5	79.4	78.0	88.5
AD_2	Duration (hr)	0.03	0.05	0.04	0.01
	Cost (USD)	$6.00	$9.00	$1.25	$0.43
	Accuracy (%)	25.0	50.0	100.0	100.0
AD_3	Duration (hr)	0.67	0.62	0.15	0.01
	Cost (USD)	$120.00	$111.00	$6.59	$3.62
	Accuracy (%)	0.0	100.0	100.0	100.0
PM_1	Duration (hr)	3.07	0.48	0.28	0.01
	Cost (USD)	$552.00	$87.00	$11.38	$6.75
	Accuracy (%)	100.0	100.0	100.0	100.0
IM_1	Duration (hr)	0.27	0.27	0.10	0.02
	Cost (USD)	$48.00	$48.00	$2.98	$3.89
	Accuracy (%)	94.0	94.0	98.0	90.0

The main criteria for document conformance measurement methods were (1) duration, (2) costs, and (3) accuracy of the results. (1) Duration included duration from executing the first module to the last module. However, initial preparation time such as software installation time was not included considering that there were many alternatives and that companies or auditors may already have the software installed, such as Toolset 3 (Adobe Pro), that can convert PDF to Word file format. Also, some software packages (e.g., Toolset 9 (Copyleaks)) were web-based, which did not require installation. Thus, duration included labor hours and software operation time (i.e., runtime). Software runtime was independent of the number of pages of documents, for both text extraction and text matching toolsets. Durations were sometimes affected by the fact that the conformance measurer may be aware of documents due to the previous cases. Notice that manual methods were sometimes quicker than the semi-automated methods (e.g., AD_2). This was an important finding because it implied that the human brain can recall previous knowledge and can utilize it for future applications.

(2) Cost included labor and software costs. For the labor costs, an average hourly rate of auditors (FERF) and workers (U.S. Bureau of Labor Statistics) in 2016 were applied because the auditor fee was based upon 2016. For the software costs, prepaid fees, such as annual or monthly fees, were not included since they were independent of the number of pages or operation time. The monthly or annual recurring fees for toolsets were stated in Chapter 5. Due to the numerous alternatives in which prices vary from open source to custom price (or price-upon-request), cost-per-use ($0.054/page) of a software package (Toolset 9 (Copyleaks)) was only considered.

(3) The accuracy in this experiment referred to the correctness of the estimates of a method against the actual. The accuracy was not the rate of conformance. It was how conformances and non-conformances were accurately detected. For example, no method detected any similarity from two documents in PM_1. Therefore, all text portions were marked as non-conformance and all methods achieved the accuracy of 100%. The accuracy is further discussed later in the section. The results of the cost breakdown are first presented in Table 6-5.

Table 6-5: Cost and Duration Breakdown

Method		DM	DS1	DS2	DS3
Unit Cost / Case ID	Labor Cost / hr	$180.00	$180.00	$25.00	$25.00
	Comparison Cost / pg	N/A	N/A	$0.05	$0.05
AD_1	Labor Hour (hr)	4.70	1.03	0.56	N/A
	Software Runtime (hr)	N/A	N/A	0.02	0.02
	Document Pages for Comparison (Module 4)	152	152	125	152
	Cost (USD)	$846.00	$186.00	$20.66	$8.21
AD_2	Labor Hour (hr)	0.03	0.05	0.03	N/A
	Software Runtime (hr)	N/A	N/A	0.01	0.01

	Document Pages for Comparison (Module 4)	8	8	8	8
	Cost (USD)	$6.00	$9.00	$1.25	$0.43
AD_3	Labor Hour (hr)	0.67	0.62	0.13	N/A
	Software Runtime (hr)	N/A	N/A	0.02	0.01
	Document Pages for Comparison (Module 4)	67	67	60	67
	Cost (USD)	$120.00	$111.00	$6.59	$3.62
PM_1	Labor Hour (hr)	3.07	0.48	0.26	N/A
	Software Runtime (hr)	N/A	N/A	0.02	0.01
	Document Pages for Comparison (Module 4)	125	125	89	125
	Cost (USD)	$552.00	$87.00	$11.38	$6.75
IM_1	Labor Hour (hr)	0.27	0.27	0.08	N/A
	Software Runtime (hr)	N/A	N/A	0.01	0.02
	Document Pages for Comparison (Module 4)	72	72	16	72
	Cost (USD)	$48.00	$48.00	$2.98	$3.89

For Table 6-5, labor hours and software runtime were rounded to two decimal places. Therefore, Table 6-5 had minor differences compared to Table 6-4 (e.g. duration). Cost depended on the number of pages of documents and labor hours. Since professional input was required for the manual method (DM) and the semi-automated method 1 (DS1), $180/hr (auditor) labor cost was applied. On the other hand, $25/hr (average worker) labor cost was applied for the semi-automated method 2 (DS2) since average worker input was enough for the information type detection module (Module 2) in DS2. Because the manual steps were skipped, the semi-automated method 3 (DS3) did not require any major human input. The numbers of pages for processing the text matching module (Module 4) were considered since the initial page numbers changed after the information type detection module (Module 2) for the semi-automated method 2 (DS2). Thus, document pages for comparison (Module 4) refers to the total pages of two documents just before executing the text matching module (Module 4). The results of Table 6-4 and Table 6-5 are further discussed in Section 6.1.3.

The accuracy was a measure to evaluate the correctness of output returned by each method. Output was an estimation of whether the text portion was in conformance. Detailed accuracy breakdown is presented in Table 6-6, Table 6-7, Table 6-8, Table 6-9, and Table 6-10. Examples are also presented to explain the errors caused by humans and machines (e.g., text matching software). The commonly accepted Equation (4) was used for the calculations. Note that the sum of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) did not include all the portions of the documents. In fact, the majority of true negative (TN) text portions were not included. This is explained further in Section 6.1.3. Assuming that the matching portions were identified by any of the methods at least once, the sum (TP, TN, FP, FN) was the text portion candidates that were estimated to be in conformance at least by one method.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad \text{(4)}$$

Where:

- TP: the number of True Positives (estimated matching & actual matching),
- TN: the number of True Negatives (estimated non-matching & actual non-matching),
- FP: the number of False Positives (estimated matching & actual non-matching), and
- FN: the number of False Negatives (estimated non-matching & actual matching)

For AD_1, the accuracy breakdown is presented in Table 6-6. Conforming portions appeared to exist substantially more than non-conforming portions. The sum of true positives (TP) and false negatives (FN) was equivalent across the methods. Also, the sum of true negatives (TN) and false positives (FP) was equivalent across the methods. This meant that for this case (AD_1), 167 text portions were matching (i.e., in conformance) while 42 text portions were not matching (i.e., not in conformance).

Table 6-6: The Accuracy Breakdown for Case AD_1

Index	Manual (DM)	Semi-Automation 1 (DS1)	Semi-Automation 2 (DS2)	Semi-Automation 3 (DS3)
TP	162	125	121	143
FP	42	1	0	0
FN	5	42	46	24

TN	0	41	42	42
TP+FP+FN+TN	209	209	209	209
Accuracy (%)	**77.5**	**79.4**	**78.0**	**88.5**

It was noticeable that the manual method (DM) had the highest true positives (TP). The high true positives (TP) implied that most conforming text portions were correctly identified by the manual method (DM). The high true positives (TP) reversely implied that the text matching software (Toolset 9 (Copyleaks) may have missed true conforming text portions. However, the results balanced out due to the manual method (DM)'s highest false positives (FP). The high false positives (FP) meant that the manual method incorrectly identified non-conforming text portions as conforming text portions. These incorrect identification caused false positives (FP) of the manual method (DM) to increase while true negatives (TN) to be at zero. On the other hand, the non-conforming text portions were never identified as conformance by the methods (DS2, DS3) that uses the text matching software (Toolset 9 (Copyleaks)).

Some examples of false negatives (FN) are listed below with explanations.

- The manual method (DM) and the semi-automated method 1 (DS1) missed matching text ("The welfare of its employees and their families") which both the semi-automated method 2 (DS2) and the semi-automated method 3 (DS3) detected. This was due to the human error. Likewise, there were five false negatives (FN) in the manual method (DM).
- The semi-automated method 1 (DS1) missed matching text ("Owners and contractors must:") which all other methods detected. This was due to manual keyword detection error. Likewise, there were 42 false negatives (FN) in the semi-automated method 1 (DS1).
- The semi-automated method 2 (DS2) missed matching text ("Provide a safe workplace") which all other methods detected. This was due to the software (Toolset 9 (Copyleaks)) error. Likewise, there were 46 false negatives (FN) in the semi-automated method 2 (DS2). This is partially due to how the software is programmed (e.g., punctuation, case sensitivity).
- The semi-automated method 2 (DS2) and the semi-automated method 3 (DS3) missed matching text ("Tamper: To alter, meddle, interfere, substitute or change") which both Manual (DM) and The semi-automated method 1 (DS1) detected. This was due to the software (Toolset 9 (Copyleaks)) error. Likewise, there were 24 false negatives (FN) in the semi-automated method

3 (DS3).

The total number of pages of AD_2 was significantly lower than AD_1. Thus, there were not as many text portion candidates as AD_1. Therefore, the accuracy gaps between the methods were greater for AD_2 than AD_1. From the accuracy breakdown in Table 6-7, it was clear that the semi-automated methods (DS2, DS3) that automated the text matching module (Module 4) had higher accuracy than the manual method (DM).

Table 6-7: The Accuracy Breakdown for Case AD_2

Index	Manual (DM)	Semi-Automation 1 (DS1)	Semi-Automation 2 (DS2)	Semi-Automation 3 (DS3)
TP	1	1	3	3
FP	1	0	0	0
FN	2	2	0	0
TN	0	1	1	1
TP+FP+FN+TN	4	4	4	4
Accuracy (%)	**25.0**	**50.0**	**100.0**	**100.0**

The semi-automated methods (DS2, DS3) that utilized the text matching software (Toolset 9 (Copyleaks)) detected all conformances correctly. The correct detection caused the true positives (TP) to be the highest and the false negatives to be at zero. Moreover, because these methods (DS2, DS3) did not identify non-conformances as conformances, the true negatives (TN) was high while the false positive (FP) was at zero.

Examples of false negative (FN) and false positive (FP) are listed below with explanations.

- The manual method (DM) and the semi-automated method 1(DS1) missed the matching text ("Intentionally left blank.") by overlooking some parts of documents. Both the semi-automated method 2 (DS2) and the semi-automated method 3 (DS3) detected this matching text portion. Likewise, there were two false negatives (FN) for the manual method and the semi-automated method 1 (DS1).
- The manual method (DM) incorrectly detected text ("In addition to the obligations set out in the Policy and this Standard, all Employees must comply with any additional site-specific Standards.") and ("In addition to the obligations set out in this Standard, all Employees must

comply with any additional Site Specific Standards.") as matching. Words, cases, and hyphen differences were missed. This is an example of false positive (FP) for the manual method.

The accuracy breakdown for AD_3 is presented in Table 6-8. AD_3 was expected to have significant amount of conformance. However, the company had modified the institution's document; thus, though similar, matching text portions were rare.

Table 6-8: The Accuracy Breakdown for Case AD_3

Index	**Manual (DM)**	**Semi-Automation 1 (DS1)**	**Semi-Automation 2 (DS2)**	**Semi-Automation 3 (DS3)**
TP	0	0	0	0
FP	2	0	0	0
FN	0	0	0	0
TN	0	2	2	2
TP+FP+FN+TN	2	2	2	2
Accuracy (%)	**0.0**	**100.0**	**100.0**	**100.0**

There were two false positives (FP) in the manual method (DM). This meant that two text portions were incorrectly detected as conformance. These incorrect detections did not occur in the semi-automated methods (DS1, DS2, DS3).

An example of the false positive (FP) is provided.

- The manual method (DM) incorrectly detected text ("Has failed to comply with the Drugs and Alcohol Work Rule,") and ("Has failed to comply with the alcohol and drug work rule,") as matching. Case sensitivity and order of text differences were missed. There were two false positives (FP).

The accuracy breakdown for PM_1 is presented in Table 6-9. Since none of the methods detected any matching portion of texts, all methods were perceived to have accurately detected non-conformances.

Table 6-9: The Accuracy Breakdown for Case PM_1

Index	**Manual (DM)**	**Semi-Automation 1 (DS1)**	**Semi-Automation 2 (DS2)**	**Semi-Automation 3 (DS3)**
TP	0	0	0	0

FP	0	0	0	0
FN	0	0	0	0
TN	0	0	0	0
TP+FP+FN+TN	0	0	0	0
Accuracy (%)	**100.0**	**100.0**	**100.0**	**100.0**

Lastly, the accuracy breakdown for IM_1 is presented in Table 6-10. The number of pages was relatively larger and text portions were alike. The breakdown result was similar to AD_1. However, the final accuracy differed from AD_1 in that the accuracy of the semi-automated method 2 (DS2) was greater than the semi-automated method 3 (DS3). In fact, DS3 scored the lowest compared to the other methods.

Table 6-10: The Accuracy Breakdown for Case IM_1

Index	Manual (DM)	Semi-Automation 1 (DS1)	Semi-Automation 2 (DS2)	Semi-Automation 3 (DS3)
TP	47	46	48	44
FP	1	0	0	0
FN	2	3	1	5
TN	0	1	1	1
TP+FP+FN+TN	50	50	50	50
Accuracy (%)	**94.0**	**94.0**	**98.0**	**90.0**

The accuracy breakdown demonstrated that there was not much of a difference across the methods. Relatively, semi-automated method 3 (DS3) had higher false negatives (FN).

An example of a false negative (FN) is provided.

- From the following matching text, “The main objective of the interface management team is to proactively identify, define, document, resolve and monitor interfaces between the identified stakeholders within a structured and traceable framework. The key elements of the interface management plan are:”, Toolset 9 (Copyleaks) failed to detect the conformance and marked “The” as non-conformance when it was actually conforming. Likewise, there were five false negatives (FN) in the semi-automated method 3 (DS3).

6.1.3 Discussion and Conclusions of the Practice Guideline Document Conformance Measurement Method

The concepts of “conformance,” “matching,” and “related” in Figure 6-2 are similar to “relevance” in the information retrieval (IR) from computer engineering field. In Section 6.1.2, accuracy was defined by applying the concept of information retrieval. Rather than having an entire text portion as the scope, by including only the relevant portions, the sensitivity of the methods was explored. Along with accuracy, the evaluation measures (precision, recall, and fall-out) from information retrieval can be used to assess the document conformance methods and may provide additional insights.

Five cases are selected for document conformance measurement (Figure 6-4). These five cases were different in terms of topics, the sources of the two documents, characteristics, and the length. The first, second, and third were related to alcohol and drug guidelines, fourth project management, and fifth interface management. There were documents from same sources such as first, second, and fifth cases. Third and fourth cases were from different sources.

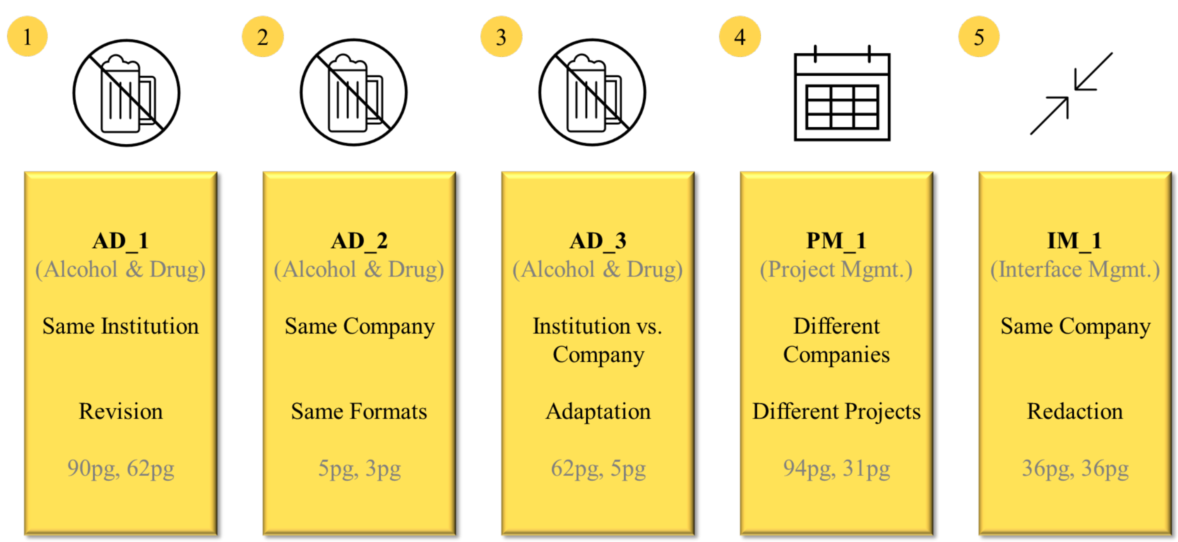


Figure 6-4: Summary of Five Representative Cases for Document Conformance Measurement

For document conformance measurement, five cases (AD_1, AD_2, AD_3, PM_1, IM_1) were experimented with four methods (DM, DS1, DS2, DS3). The results validate that the semi-automated methods (DS1, DS2, DS3) were quicker, less expensive, and more accurate than the manual method (DM) in general (Table 6-4). Other implications and findings are elaborated below.

The first case (AD_1) shows the strength of semi-automation in terms of duration and cost. The two documents that were used for the first case had the largest number of pages in total among all cases. These two documents from AD_1 took the longest hours in manual checking not only because it had the largest number of pages in total but because the documents had both similarities and differences. The non-conformance detection was not as simple as other cases (e.g., AD_2, PM_1). As the automation level increased, the duration and the cost substantially decreased up to 200 times for the duration reduction and 100 times for the cost reduction. When it comes to accuracy, the semi-automated method 3 (DS3) reached the highest accuracy. Considering that DS3 had the least manual tasks, it was a positive outcome that can save time and costs with higher accuracy.

Figure 6-5 is the results of AD_1. X-axis represents automation level. Y-axis represents cost, duration, and error rate. The expectation is that the cost, duration, and error rate to be at the bottom. The chart shows that compared to the manual method, semi-automated method 3 is not only more accurate but takes less time and costs. The result is because machines do not get tired or have emotions while humans can misread or simply miss things. AD_1 was an example of the longest documents. This gap of cost, duration, and error rate (=100 - accuracy (%)) may increase when the documents are even longer.

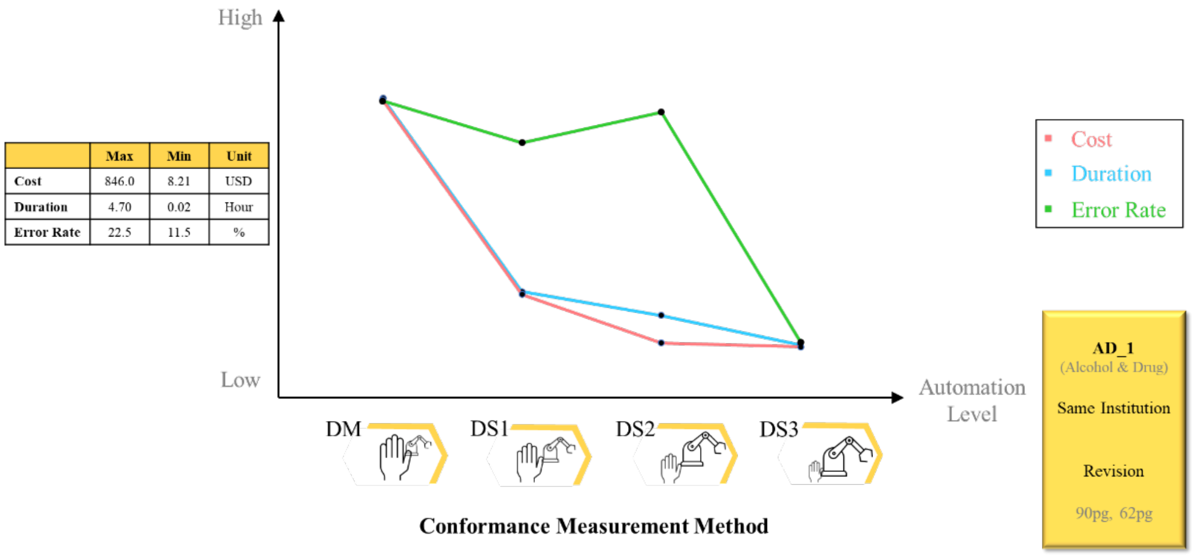


Figure 6-5: The Results of AD_1

The second case (AD_2) compares how a small amount of documents can impact accuracy. Duration among the methods (DM, DS1, DS2, DS3) was not much of a difference, but accuracy was greater in high-level automation (DS2, DS3). This result can be attributed to the difference between humans and the machine (or software). Though the machine did not detect every non-conformance correctly, in this second case (AD_2), the text matching software (Toolset 9 (Copyleaks)) detected what the manual method (DM) had missed.

The third case (AD_3) illustrates what happens when a company adopts a guideline from an institution or another company. Similar to AD_1, because both similarities and differences existed, manual method (DM) took longer than the other methods (DS1, DS2, DS3). The duration was measured with integrity, but the measurements were not in proportion to pages. Some cases were quicker than others due to the font size, inclusion of images, or even due to the preview experiences. For instance, in AD_1 and AD_3, the same document (one of two documents) was used. Because AD_1 was conducted before AD_3, the fact that the document had been previewed, influenced the timing for AD_3. When it comes to accuracy, the text matching software (Toolset 9 (Copyleaks)) detected what the manual method (DM) had missed.

Preview experiences also affected the duration within a case. For the fourth case (PM_1), the duration difference between the manual method (DM) and the semi-automated method 1 (DS1) was large because semi-automated method 1 (DS1) was affected by previous conformance measurement (DM). However, semi-automated method 3 (DS3) clearly took the least amount of time to conclude that no text portion is in conformance (or matching). Considering that the semi-automated method 3 (DS3) was not affected by the other methods that involve manual tasks (DM, DS1, DS2), it can be concluded that DS3 is the most efficient method.

The difference between the semi-automated method 2 (DS2) and the semi-automated method 3 (DS3) is that the semi-automated method 2 (DS2) includes Module 2 (information type detection). Module 2 is manually done and typically increases time and costs, but for IM_1 case, the semi-automated method 2 (DS2) reduced cost and error rate than the semi-automated method 3 (DS3) (Figure 6-6). The cost includes labor hours and the length of a document. For IM_1 case, only part of the documents were needed to be compared. The conclusion about DS3 being the most efficient method among other methods did not change in IM_1. However, when it comes to the costs and the accuracy, the semi-automated method 2 (DS2) performed better than the semi-automated method 3 (DS3). This is perceived to be due to the information type detection module (Module 2). Module 2 not only reduced the total number of pages to reduce the software cost but removed unnecessary tables or images that may hinder appropriate conformance measurement. As a result, higher accuracy was reached. However, in AD_1, the main reason for the underperformance of the semi-automated method 2 (DS2) was because of the text extraction module (Module 1) that caused the format to change from .pdf to .docx. Significant information losses occurred resulting in lower accuracy when completing the information type detection module (Module 2).

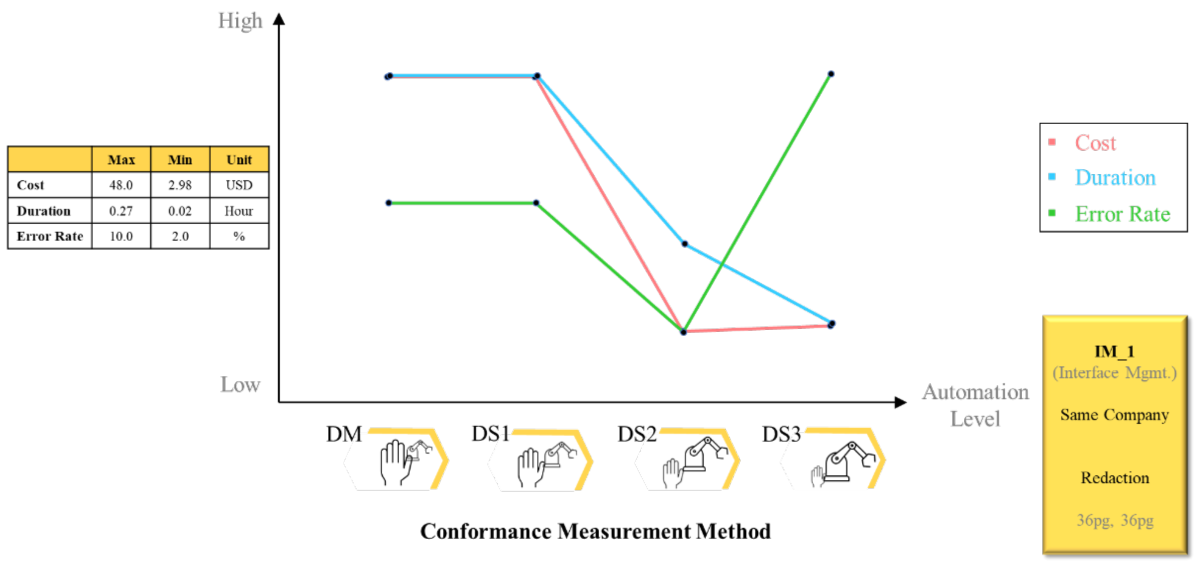


Figure 6-6: The Results of IM_1

Likewise, every case had implications. The implications differed due to the relationships between two documents, the numbers of pages of documents, and the topics of cases. Some fluctuations in the assumptions of the unit costs, such as labor cost-per-hour and/or comparison cost-per-page may change the results.

The difference between the currently used average auditor cost ($180/hr) and average worker cost ($25/hr) is over seven times (Table 6-5). Unless the initial fee for the software package outweighs the benefits of semi-automation, the semi-automated methods are recommended for practice guideline document conformance measurement. In fact, the comparison cost-per-page varied from $0.05 to $0.1 in Toolset 9 (Copyleaks) depending on the plan that an auditor or a company may purchase. However, the variance of the cost-per-page is perceived not to change the overall results.

Labor hour and software runtime were distinguished under the assumption that software operation does not require any human input. Thus, software runtime did not affect the cost. The software runtime included the text extraction module (Module 1) and the text matching module (Module 4). The results demonstrated that the number of pages or the degree of conformance did not highly affect the timing for the semi-automated methods (DS2, DS3).

Among the semi-automated methods, the semi-automated method 3 (DS3) was promising, provided that the semi-automated method 2 (DS2) did not prove substantially beneficial (e.g. being unable to reduce the large number of pages). For the semi-automated method 2 (DS2), labor hours was equivalent to the manual input for the information type detection module (Module 2). Since this module (Module 2) can be completed by non-professionals, there were advantages in costs. However, Module 2 increased duration without any accuracy benefit with the exception of IM_1. For the semi-automated method 3 (DS3), labor hours were ignored. In reality, since software packages were not integrated, minimum human intervention was unavoidable. The software runtime was not included for the manual method (DM) and the semi-automated method (DS1), since additional software (other than Word or PDF viewer) was not necessary.

When it comes to accuracy, if all true negative (TN) text portions were included in the accuracy equation, all the accuracy results would have increased. Not because there were too many text portions, but because too many portions will reduce the accuracy gap greatly, to clearly differentiate the results among the methods, text portion candidates were selected for the results. The text portion candidates were the text portions that at least one method detected as conformance. Thus, if all methods identified as non-conformance, the text portion was not considered as a candidate.

The text portion is defined as a part of body of texts. Text portions must be distinguished from phrase, sentence, or paragraph since documents had mixed of all. For example, documents used in AD_1 included lists in which some were phrases and others were group of sentences. Newline, indentations, and numberings were some ways to determine portions. While there were ambiguous situations such as one document has period and the other not, consistent rules were applied to accurately measure the conformance.

Examples of false positives (FP) or false negatives (FN) from the manual method (DM) were mostly due to the punctuation or conjunction differences. The reason behind the false negatives (FN) in the semi-automated methods (DS2, DS3) is not certain. However, considering that the software uses OCR, it may be due to the wrong recognition of letters or ways of partitioning the body of texts. Moreover, how the software is programmed may affect the results. In spite of these errors, the semi-automated methods were consistently more accurate than the manual. Moreover, as the automation level increased, the duration of operation decreased substantially.

6.2 Validation of the Benchmark Workflow Conformance Measurement Method

For the validation of the three methods (WM, WS1, and WS2), four workflow instances were selected out of the 598 workflow instances. When the sample size is expanded, the conformance measurement duration increased at different rates. This is because while the duration of WM and WS1 increases proportionally to the number of workflow instances, the duration of WS2 increases proportionally to the number of flow paths. Each workflow instance has around 10-30 flow paths. Because there were overlapping flow paths as the workflow instances were grouped in WS2, the number of flow paths in WS2 was significantly lower than WS1. Thus, the duration for measuring conformance was reduced. In summary, the semi-automated method 2 (WS2) is the recommendation for a large number of workflow instances considering duration. It is hard to generalize accuracy when the number of workflow instances increase because all three methods have potential for human errors. Additionally, discussion to validate the application of performance measurement module (Module 2) as well as results with different conformance definitions are included in Section 6.2.3.

6.2.1 Implementation of the Benchmark Workflow Conformance Measurement Method

For the validation, the practice benchmark workflow conformance measurement flowchart was applied (Figure 4-7) and the functional requirements were considered (Figure 4-4). To validate the workflow conformance measurement methods, in this study, an automated change request process data from a capital project was collected.

The change request process is an essential process that is included in most capital projects, because changing original plans or workflows is sometimes inevitable. An effective change request workflow can reduce the overall duration to complete the project, and therefore has significance to the improvement and maintenance of well-defined and effective workflows. These workflows are created by a company in which the workflow operates, such as an architecture, engineering and construction (AEC) firm. Furthermore, in order to achieve higher conformance and productivity, some companies develop automated workflow engines (event driven software stacks running on a server or in a cloud-based computing environment).

Figure 6-7 is the workflow engine's graphical programing interface and it served as an illustration of the Workflow implementation 1. A benchmark workflow was designed based on Workflow implementation 1 with some changes. The workflow implementation includes all activities and flow paths (i.e., connections) that were allowed in the change request process, by both humans and the workflow engine. All electronic workflow implementations that were driven by the automated workflow engine are provided in Appendix F. These workflow implementations were previously studied by Karimidorabati (2014). A workflow implementation refers to an implemented workflow for a specific organization or project while workflow instance refers to an executed instance of an implemented workflow (Golzarpoor, 2016). A workflow implementation is created for specific projects from a workflow template, a customized workflow with the most common components, such as IFP: Industry Foundation Process (Golzarpoor, 2016).

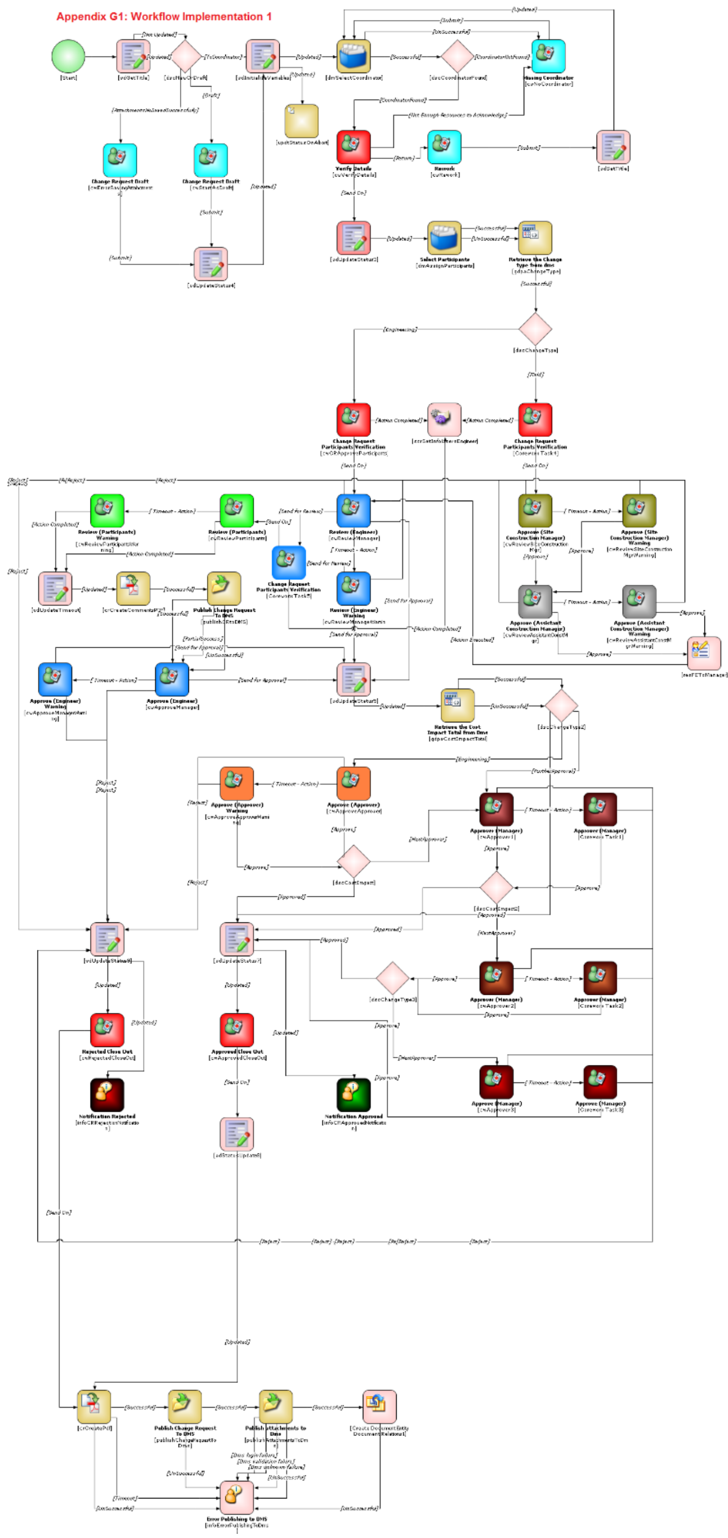


Figure 6-7: Automated Workflow Engine Representation (Workflow Implementation 1) of a Change Request Process

However, a benchmark workflow differs from a workflow implementation (Figure 6-8). The benchmark workflow (Figure 6-9) refers to a desired or pursued workflow for conformance measurement, unlike workflow implementation which is the workflow that is implemented for process automation. The purpose of workflow implementation is to improve the workflow for better performance. By adding events, flow paths, documents and decisions, workflow can be more streamlined and effective. An automated workflow engine has enforcement to conform to the workflow implementation while benchmark workflows do not have this enforcement. On the other hand, the purpose of benchmark workflow is to measure conformance. Measuring can demonstrate how and why to improve the workflow implementation.

A workflow implementation only allows events and flow paths that are predefined. However, sometimes deviations from the workflow implementation may bring better performance. With having the business adjust the workflow implementation based on real-world findings, a benchmark workflow can expedite findings and reveal flaws within the implementation. Benchmark workflow can also assist in determining future adjustments. The benchmark workflow allows the business to be proactive rather than reactive.

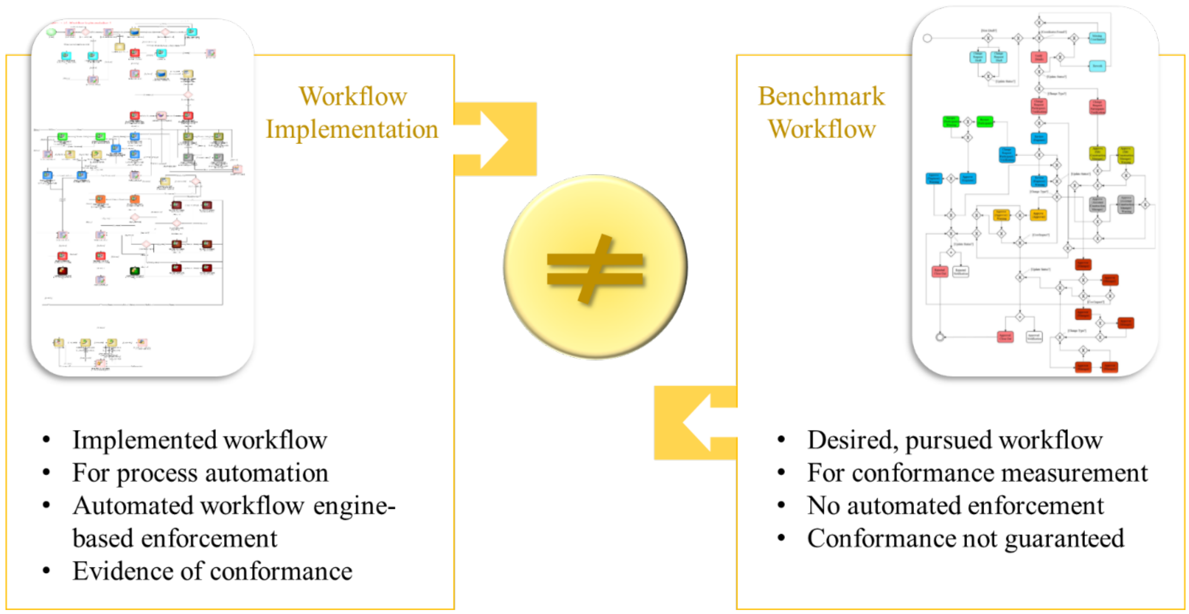


Figure 6-8: The Differences between Workflow Implementation and Benchmark Workflow

In Figure 6-7, human-involved activities are represented by person icons on the top left of the activity boxes. Examples of human-involved activities include: “Change Request Draft,” “Verify Details,” “Change Request Participant Verification,” “Review,” “Approve,” “Warning,” “Notification,” and “Close Out.”

It is also worthwhile to note the colours of the boxes (i.e., activities). Boxes of the same colour represent that the same department or team is responsible or affected by the associated activities. For example, “Change Request Draft,” “Rework,” and “Missing Coordinator” in the light blue colour were executed by the same department.

Pink boxes indicate activities automatically completed by the workflow engine and without human involvement. These activities include: “Get Title,” “Initialize Variables,” “Update Status,” “Select Participants,” “Retrieve Change Type from DMS (Document Management System),” “Set Info Users,” “Update Timeout,” “Create Comments PDF,” “Publish Change Request to DMS,” “Retrieve the Cost Impact Total from DMS,” “Publish Attachment to DMS,” “Create Document Entity (Document

Relations),” and “Error Publishing to DMS.” Pink diamonds represent yes or no queries, such as “New Draft (new or not),” “Coordinator Found (found or not),” “Change Type (engineering or field),” and “Cost Impact (approved or not).” Every case differed according to the responses to the queries.

Not many changes were made throughout the workflow implementations; however, any changes made are indicated in workflow engine activities (pink boxes) in Appendix F. For example, workflow implementation 2 adds a workflow engine activity called “Update Status in Review,” affecting the benchmark workflow by allowing transitions from “Approver (Manager)” to “Review (Engineer).” As versions were updated, some activities were added while others were removed. Most changes only affected workflow engine operations. However, the change from workflow implementation 1 to 2 is a great example of how workflow implementation changes over time.

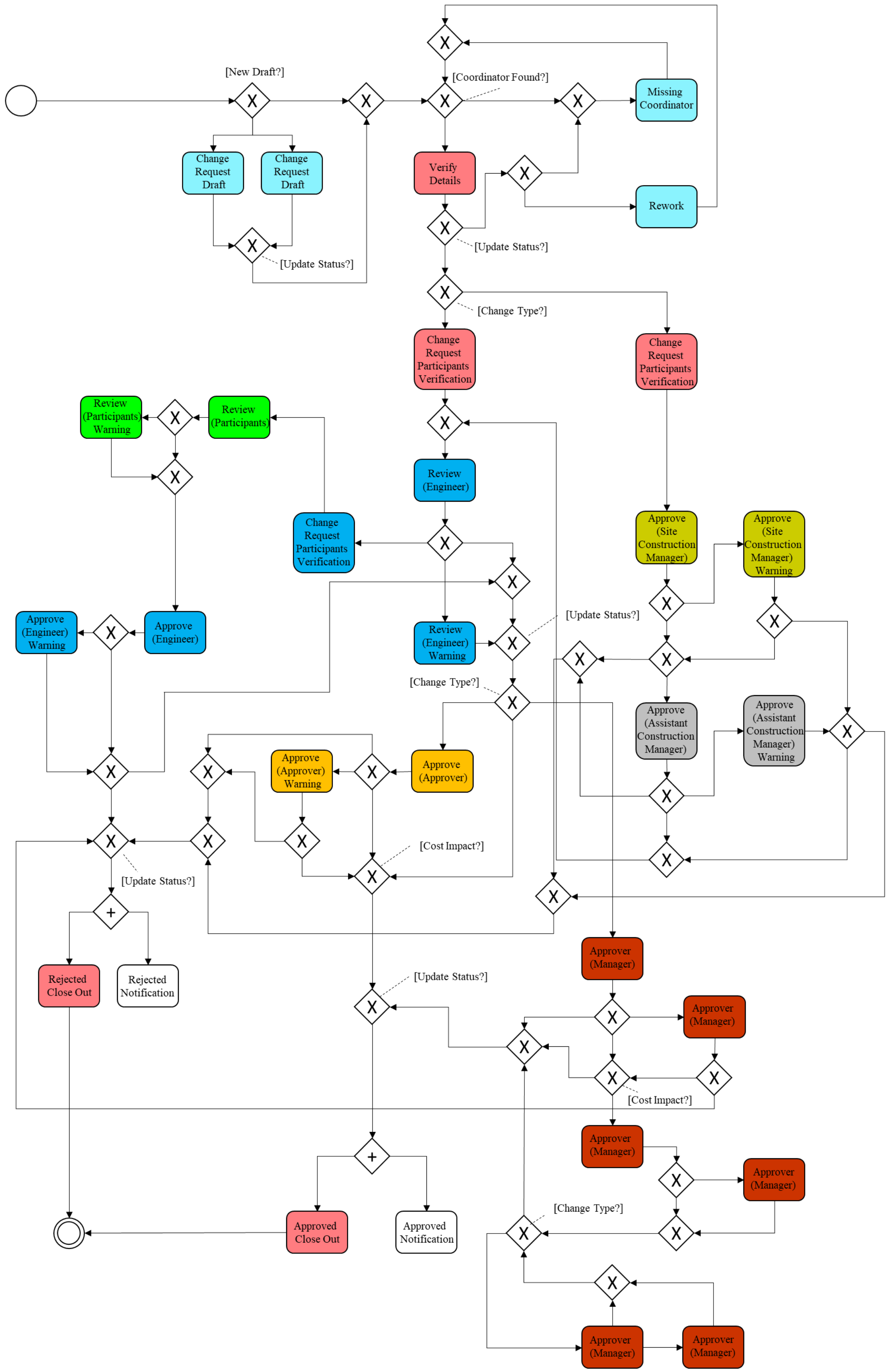


Figure 6-9: The BPMN-Format Benchmark Workflow (Original Benchmark Workflow)

If a process is followed with workflow engine representation, conformance can, to some extent, be enforced in automated ways. If the process is clearly defined in the form of a benchmark workflow, an event log can be compared against the benchmark workflow to measure conformance. The event log results from the execution of the process. The event log typically includes the case IDs, activities, and timestamps. The change request event log data and eight versions of workflows were collected from an oil and gas capital project. The event log contained a six-month record of a change request process in total of 598 workflow instances and 16,633 activities. The event log consisted of four types of change requests: the vendor change request (VCR), engineering change request (ECR), field change request (FCR), and contract change request (CCR). All four types followed the same overall process as they journey from initiation to approval or rejection. However, every workflow instance was unique because each involved distinctive activities, transitions, sequences, and timestamps. The attained event log had twelve columns: "Change_Type," "WF_ID," "Doc_ID," "Activity_Display_Name," "Created_Date_Time," "Ownership_Date_Time," "Completed_Date_ Time," "Response_by," "Name," "Current_Staus," "Process_Name," and "Version." The definitions and examples are listed in Table 6-11.

Table 6-11: Definitions and Examples of Event Log Items

Event Log Item (Spreadsheet Column)	Definition	Examples
Change_Type	Type of change	CCR, ECR, FCR, VCR
WF_ID	Workflow instance ID; Case ID	26, 27, 28
Doc_ID	Document ID	ABC-ECR-S-00026-0008
Activity_Display_Name	Name of the activity	Verify Details, Review (Engineer), etc.
Created_Date_Time	Activity created time	3/3/2020 0:34
Ownership_Date_Time	Activity owned time	3/3/2020 0:38 or NULL
Completed_Date_Time	Activity completed time	3/10/2020 1:31 or NULL
Response_by	Deadline for the activity	3/6/2020 23:00 or NULL
Name	Name of the Participant	Amy
Current_Status	Status of the activity	Abort, Approve, Closed, Submit, Reject, etc.
Process_Name	Name of the entire process	Change Request.111222333
Version	Version of the benchmark workflow that is followed	1, 2, 3, 4, 5, 6, 7, 8

"Change_Type" refers to the four types of changes. "WF_ID" referred to each workflow instance; thus, it was used as a case ID. "Doc_ID" can also be the case ID; however, it appeared more complicated than "WF_ID." "Activity_Display_Name" referred to the activity that was executed. "Created_Date_Time" was the start time of activity. "Ownership_Date_Time" referred to the time that a participant started the action. "Completed_Date_Time" was the time when the participant completed the activity. "Response_by" referred to the deadline. "Name" was the participant who owned the activity. "Current_Status" was the status of the activity. "Process_Name" was the name of this entire process. Lastly, "Version" indicated which one of the eight versions of the workflows that was followed. Among the information, "WF_ID," "Activity_Display_Name," "Created_Date_Time," "Completed_Date_ Time," and "Response_by" were specifically used for the application experiment.

For the experiment, a manual method (WM) and two the semi-automated methods (WS1, WS2) were designed to test the accuracy and compare the durations. Three methods are listed in Table 6-12.

Table 6-12: Method Descriptions for the Application Experiment

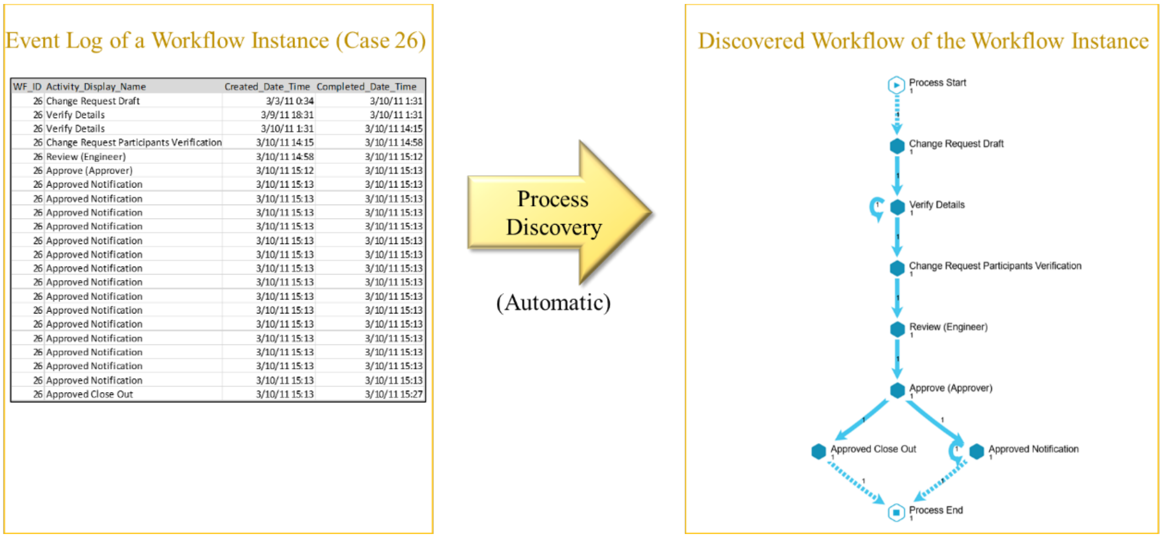
Method	Automation Level	Selected Toolsets	Sequence of Modules	Details
Manual (WM)	Level 0	• N/A	Module 3 (Manual)	• Manually compare the event log to the benchmark workflow (Log-to-Model)
Semi-Automation (WS1)	Level 1	• Toolset 14 (Celonis)	Module 1 (Semi) → Module 2 (Auto) →Module 3 (Manual)	• Semi-automatically replace the event log "NULL" timestamps • Automatically utilize the process discovery function to create a discovered workflow of individual workflow instances • Manually compare the discovered workflows to the benchmark workflow (Model-to-Model)

Semi-Automation (WS2)	Level 1	• Toolset 14 (Celonis)	Module 1 (Semi) →Module 2 (Auto) →Module 3 (Manual)	• Semi-automatically replace the event log "NULL" timestamps • Automatically utilize the process discovery function to create a discovered workflow of a group of workflow instances • Manually compare the discovered workflow of a group of workflow instances to the benchmark workflow (Model-to-Model)

*Module 1: Event Log Preprocessing; Module 2: Performance Measurement; Module 3: Conformance Measurement; Module 4: Workflow Conformance Map

Based on the flowchart (Figure 4-7), first, an event log and a benchmark workflow were selected. Then, it was decided that among the three methods, the first method be the manual method (WM) comparing the event log to the benchmark workflow. Because the manual method (WM) did not require any toolset, the event log preprocessing module (Module 1) was not necessary. The performance measurement module (Module 2) was skipped. Next, the conformance was defined. To align with the other methods, for the manual method (WM), the conformance was defined as non-existence of non-conformance. Non-conformance was defined as existence of activities or flow paths in the event log that did not exist in the benchmark workflow. Thus, NULL timestamps in the event log were considered non-conformances. When the conformance measurement occurred, the benchmark workflow and the workflow instances from the event log were compared (a log-to-model comparison).

The second method (WS1) utilized individually discovered workflows from workflow instances (Figure 6-10). On the left is the event log of a workflow instance or case and with the process discovery software, the event log can be transformed into a discovered workflow shown on the right.



WF_ID	Activity_Display_Name	Created_Date_Time	Completed_Date_Time
26	Change Request Draft	3/3/11 0:34	3/10/11 1:31
26	Verify Details	3/9/11 18:31	3/10/11 1:31
26	Verify Details	3/10/11 1:31	3/10/11 14:15
26	Change Request Participants Verification	3/10/11 14:15	3/10/11 14:58
26	Review (Engineer)	3/10/11 14:58	3/10/11 15:12
26	Approve (Approver)	3/10/11 15:12	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Notification	3/10/11 15:13	3/10/11 15:13
26	Approved Close Out	3/10/11 15:13	3/10/11 15:27

Figure 6-10: Automatic Process Discovery with a Process Mining Software (Celonis)

Since the semi-automated method 1 (WS1) required software assistance (process discovery function), the event log preprocessing module (Module 1) was necessary. The KPIs definition step from the performance measurement module (Module 2) was skipped. However, a part of the performance measurement module (Module 2) was applied. Using the process discovery software package (Toolset 14 (Celonis)), four discovered workflows were created. The discovered workflow referred to the workflow that actually occurred after an execution of a process. The conformance was defined the same as the manual method (WM) except for the NULL timestamp-related definition since NULL timestamps did not apply due to the event log preprocessing module (Module 1).

In the conformance measurement module (Module 3), the discovered workflow can be compared against the benchmark workflow for conformance measurement as shown in the purple highlights (Figure 6-11). There are two non-conformances that are identified from the comparison. Conformance measurement is achieved manually.

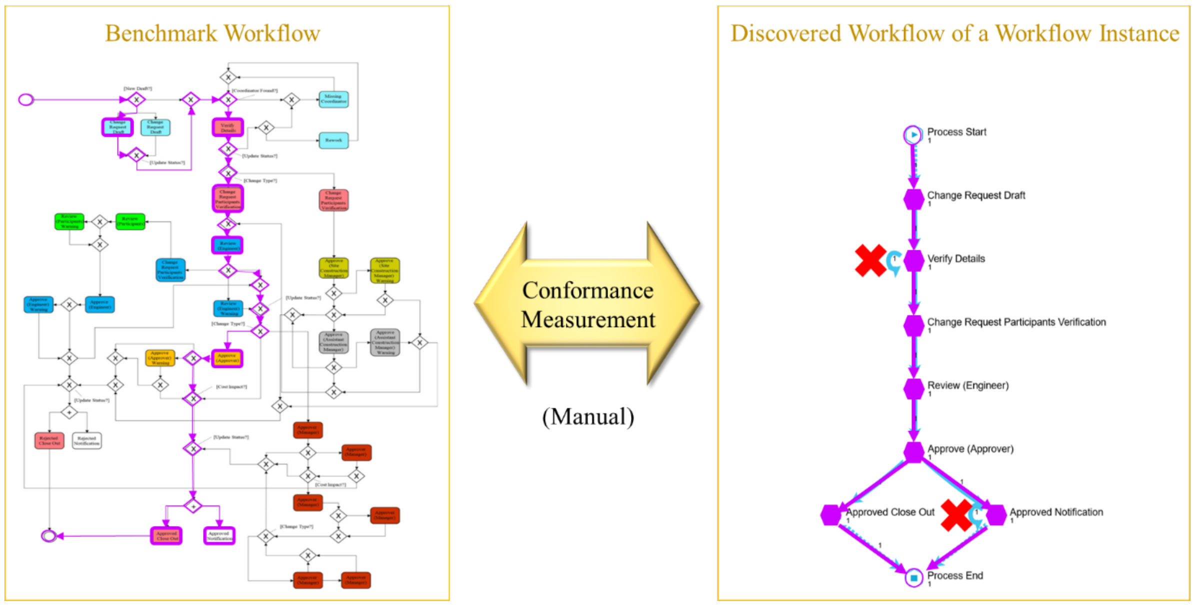


Figure 6-11: Manual Conformance Measurement Using a Discovered Workflow

The third method (WS2) also utilized process discovery function. However, instead of creating individually discovered workflows, discovered workflows from all workflow instances was integrated as a single workflow. The semi-automated method 2 (WS2) applied the same modules as the semi-automated method 1 (WS1). Strictly speaking, since KPIs are not defined and performance is not measured, groups of workflow instances cannot be created (Module 2). However, for this method validation experiment, workflow instances are grouped in the semi-automated method 2 (WS2).

Figure 6-12 is a screenshot of an event log, process mining software, and a benchmark workflow.

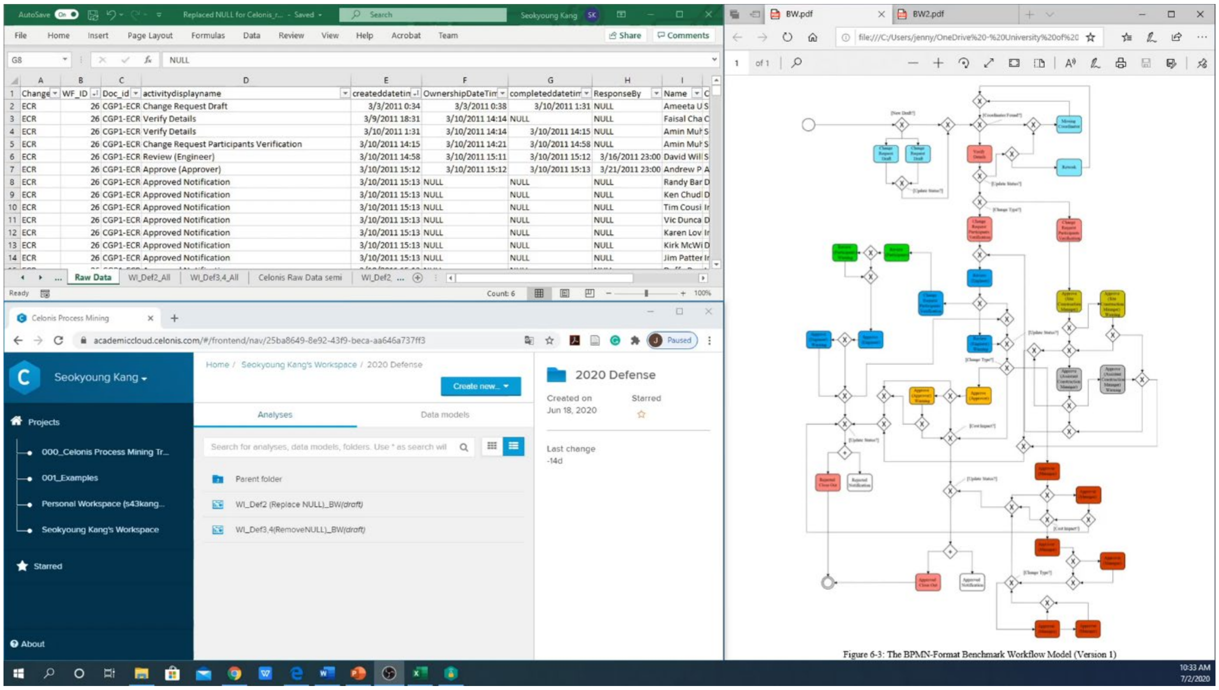


Figure 6-12: A Screenshot of Event Log (Top Left), Process Mining Software (Bottom Left), and Benchmark Workflow (Right)

6.2.2 The Results and Analysis of the Benchmark Workflow Conformance Measurement Method

The objective of this validation experiment was to compare the efficiency, cost-effectiveness, and correctness of each method in terms of their duration, cost, and accuracy. The benchmark workflow (Figure 6-9) was used to compare against the event log (WM) or the discovered workflows (WS1, WS2). The duration, cost, and accuracy were defined and measured. Duration included labor hours and software operation time but did not include the initial preparation/setup time such as creating the separate benchmark workflow in BPMN format or devising ways to preprocess the event log. Cost included labor cost but excluded software costs because their price was only available upon request. For the labor cost, a 2016 average hourly rate for auditors (FERF) and workers (U.S. Bureau of Labor Statistics) were applied. The accuracy referred to as the correctness in detecting the non-conformances. Table 6-13 summarizes the results measuring duration and cost for 598 workflow instances.

Table 6-13: Summary of 598 Workflow Instances Using Workflow Conformance Measurement Methods

Criteria	Manual (WM)	Semi-Automation 1 (WS1)	Semi-Automation 2 (WS2)
	Level 0	Level 1	Level 1
Duration (hr)	12.95	11.21	0.62
Cost (USD)	$323.75	$280.25	$15.50

The manual method (WM) took 12.95 hours to process 598 workflow instances. The difference between the manual method (WM) and the semi-automated methods (WS1, WS2) was how the executed processes were presented. While the manual method (WM) had a list of rows describing the activities, the semi-automated methods (WS1, WS2) presented visual workflows of activities. Thus, although the manual method (WM) was more specific and included more information, the semi-automated methods (WS1, WS2) were more intuitive and visual. Therefore, it took a shorter time for the semi-automated methods (WS1, WS2) than the manual method (WM). Furthermore, the semi-automated method 2 (WS2) took less than the semi-automated method 1 (WS1). This was because WS2's workflow included all workflow instances combined into one flowchart as a group, which eliminated the time spent on repetitive processes, such as identifying the same non-conformance type for more than one workflow instance.

If an auditor desires to know which non-conformance types each workflow instance has, WM or WS1 must be used. In fact, if an auditor desires to know when and how many times non-conformances exists, WM may be the only option. Most times, however, acknowledging the types of non-conformances is enough for the audit to improve the workflow implementation for the future.

The semi-automated methods (WS1 & WS2) required the event log preprocessing module (Module 1). However, the duration of this module was not included in the duration (Table 6-13) since the calculation depends on the definition of conformance. One way to deal with NULL timestamps is to remove all NULL timestamps. With excel formula, this will take less than a minute. When the NULL timestamps need to be replaced, it is more complicated. If there is a logic, it can be reflected in the excel formula, but if not, the timestamps need to be replaced manually which will take longer duration.

Process discovery functionality was utilized for the semi-automated methods (WS1 & WS2) in the performance measurement module (Module 2). Four out of the 598 workflow instances are transformed into discovered workflows for semi-automated measurements. Each workflow instance becomes an individually discovered workflow in semi-automation 1 in Figure 6-13.

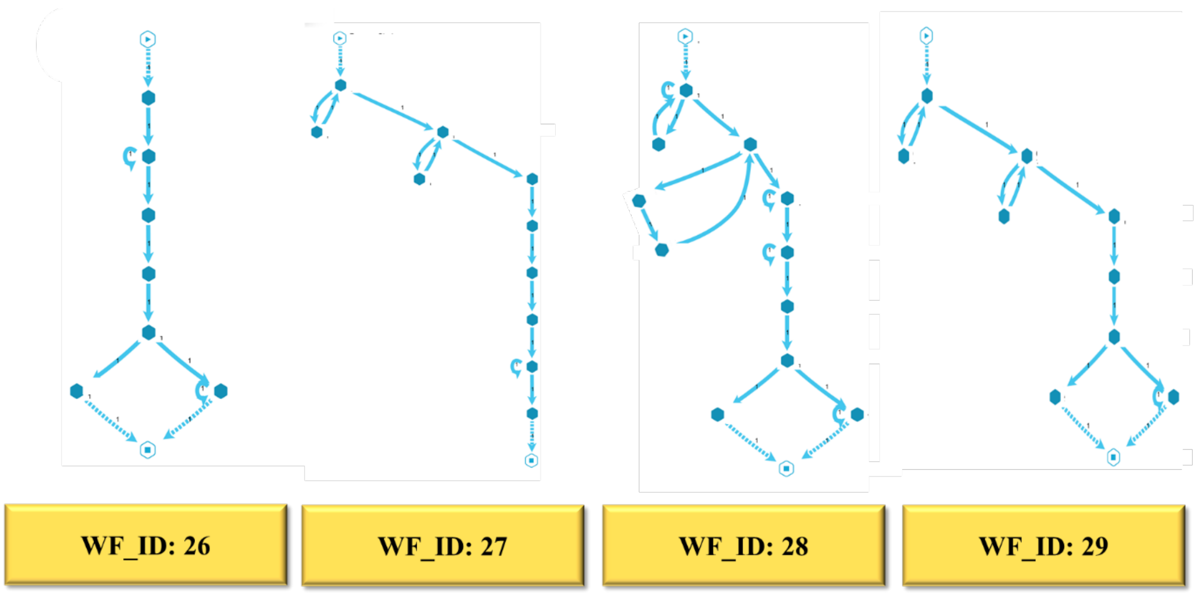


Figure 6-13: The Individually Discovered Workflows (WS1) (Source: Celonis)

All four workflow instances become a single discovered workflow in semi-automation 2 in Figure 6-14.

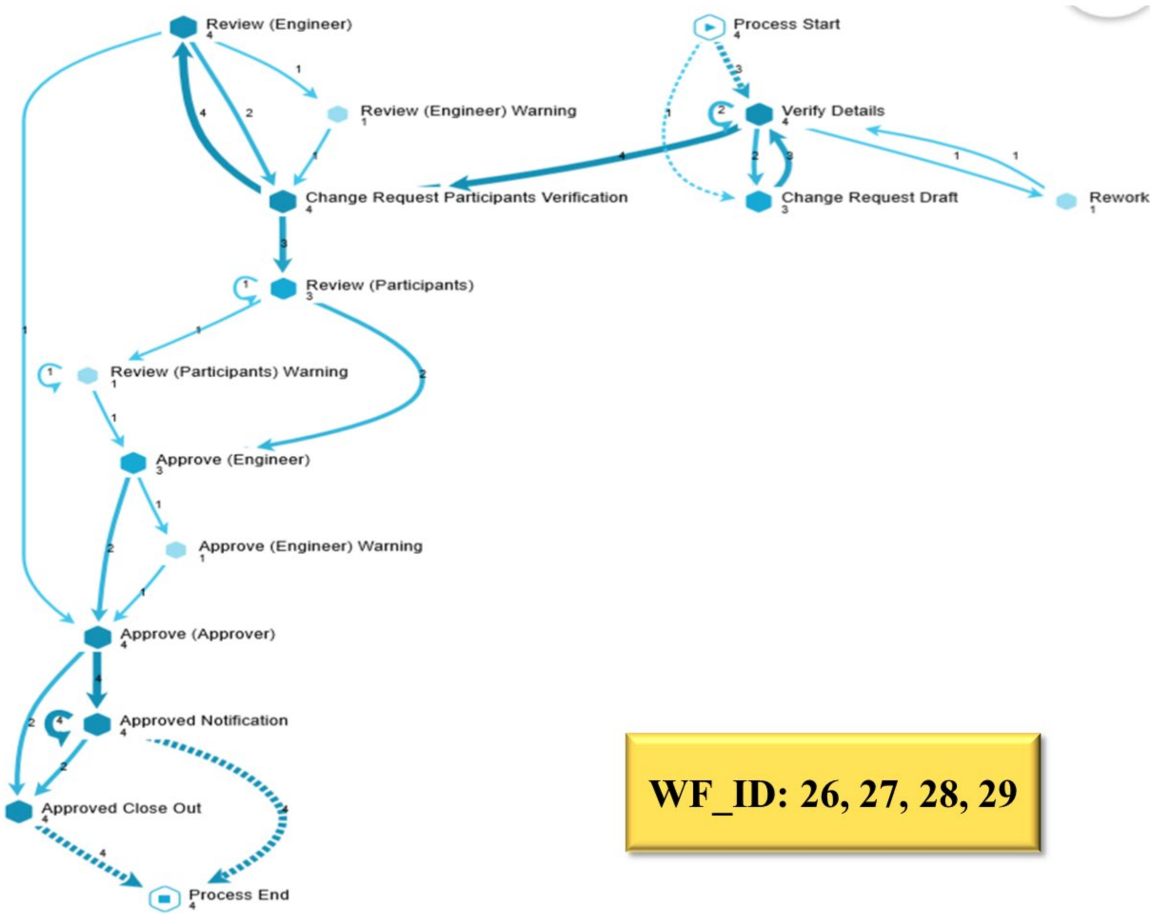
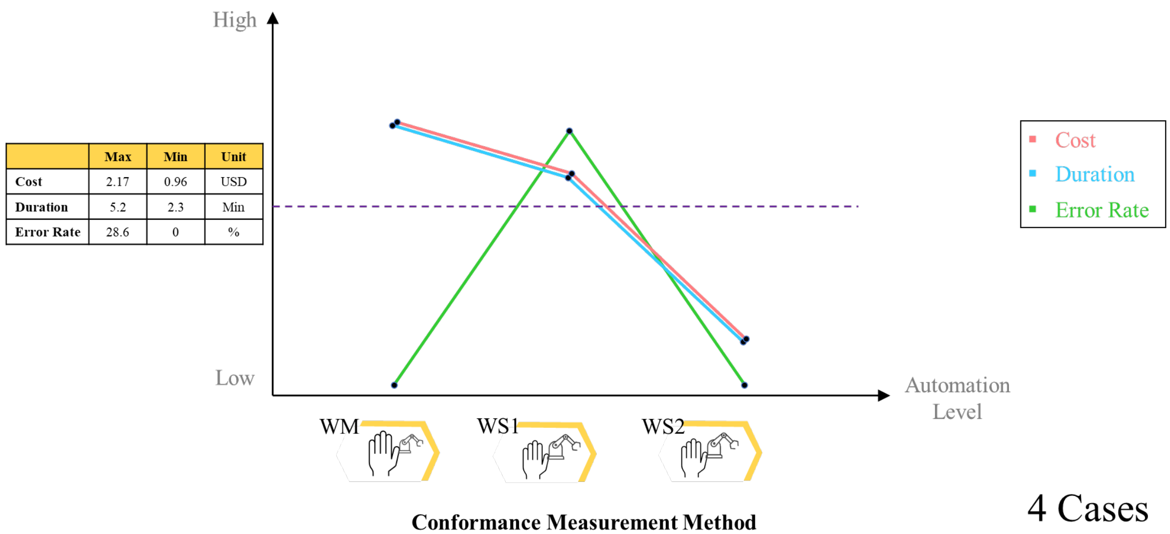


Figure 6-14: The Single Discovered Workflow (WS2) (Source: Celonis)

To upload the data and to run the event log, the process mining software (Toolset 14 (Celonis)) required the data model and analysis naming and the timestamp formatting. Altogether, with some manual input, it took approximately one minutes to run the software. Software processing time was a one-time occurrence and did not increase time proportionally to the number of workflow instances; thus, was not included in the duration. Cost was proportional to duration because the average worker's hourly rates were applied for all three methods ($25.00/hr). The task was simple and repetitive that average worker can do.

In Figure 6-15, the manual method (WM) is compared against semi-automated methods 1 and 2 (WS1 & WS2). Semi-automated method 2 (WS2) shows the decrease of more than half the cost and duration compared to the manual method. However, depending on the definition of error rate, the gap between the highest and lowest error rate will change. When scaled out over 598 cases, the error rate is outweighed by the benefits of cost and durations.



	Max	Min	Unit
Cost	2.17	0.96	USD
Duration	5.2	2.3	Min
Error Rate	28.6	0	%

Figure 6-15: The Results of Four Workflow Instances (26, 27, 28, 29)

An accuracy (=100 - error rate (%)) is a measure used to evaluate each method and can be defined in more than one way. For instance, when a flow path is non-conforming from a single workflow instance, the accuracy can be one over all flow paths of the workflow instance. However, if second workflow instance has the same non-conformance, depending on the definition of accuracy, non-conformance can be one (ignore the overlapping non-conformance) or two (count both non-conformances) over all flow paths (path types from two workflow instances or sum of all paths). As various definitions of accuracy can exist, it was perceived that exploring the methods in terms of accuracy to demonstrate that quantitative validation is possible. An example to apply the definition of accuracy is presented with four selected workflow instances. For this example, the accuracy increased when the non-conformance was detected correctly, and decreased when non-conformance was not detected or conformance was detected as non-conformance. The definition ignored the correctly detected conformance for all three methods to increase the precision of the methods. Overlapping non-conformances were counted as one.

A summary of accuracy breakdown is presented in Table 6-14. Equation (5) was used for the calculations. Note that the sum of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) did not include every flow path of the workflow. In fact, it did not include any activity (This was because there were no skipped or inserted activity). Also, the definitions of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) differ from Equation (4). Assuming that the non-conformance was identified by any of the methods at least once, the sum referred to as the potential candidates that may be in non-conformance.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (5)$$

Where:

- TP: the number of True Positives (estimated non-conformance & actual non-conformance),
- TN: the number of True Negatives (estimated conformance & actual conformance),
- FP: the number of False Positives (estimated non-conformance & actual conformance), and
- FN: the number of False Negatives (estimated conformance & actual non-conformance)

Table 6-14: The Accuracy Breakdown for Workflow Instances 26, 27, 28, and 29

Index	**Manual (WM)**	**Semi-Automation 1 (WS1)**	**Semi-Automation 2 (WS2)**
TP	7	5	7
FP	0	0	0
FN	0	2	0
TN	0	0	0
TP+FP+FN+TN	7	7	7
Accuracy (%)	**100.0**	**71.4**	**100.0**

Seven non-conformance types were detected from this example definition of accuracy. With the manual method (WM) and semi-automated method 2 (WS2), all the non-conformance types were correctly identified. On the other hand, with the semi-automated method 1 (WS1), not all the non-conformance types were correctly identified. Human errors occurred. It must be noted that all three methods required manual work for the core task (conformance measurement module (Module 3)). Thus, as much as the manual method (WM), the semi-automated methods (WS1 & WS2) had potential for human errors. These four workflow instance experiments cannot, therefore, represent accuracy of 598 workflow instances or generalize for extended circumstances. Table 6-15 presents in more detail the non-conformances that were detected from the methods.

Table 6-15: A Summary of Non-Conformance Detection and Accuracy

Type	Non-Conformances	WM	WS1	WS2	Actual
1	Verify Details→ Verify Details	X	X	X	T
2	Verify Details → Change Request Draft	X	X	X	T
3	Review (Engineer) Warning → Change Request Participants Verification	X	X	X	T
4	Review (Participants) → Review (Participants)	X	-	X	T
5	Review (Participants) Warning → Review (Participants) Warning	X	-	X	T
6	Approved Notification → Approved Notification	X	X	X	T
7	Approved Notification → Approved Close Out	X	X	X	T
$Accuracy\ (\%) = \frac{TP + TN}{TP + TN + FP + FN}$		100.0	71.4	100.0	N/A

*The X mark represents that the method detected non-conformance.
**For Actual, the T mark represents that the non-conformance actually existed and the F mark represents that the non-conformance actually did not exist.

The error from the semi-automated method 1 (WS1) was an error of conformance measurer overlooking the discovered workflow where non-conformance existed (FN).

6.2.3 Discussion and Conclusions of the Benchmark Workflow Conformance Measurement Method

When it comes to duration comparison, the manual method (WM) took longer than the semi-automated methods (WS1 &WS2), due to the time spent searching for non-conformance within the event log. When the cases became more complex, the duration for the manual method (WM) is expected to increase. On the other hand, because the discovered workflows of the semi-automated methods (WS1 & WS2) were in the form of an image, it may be easier to spot the differences compared to the manual method (WM). Thus, when the cases grow, it is expected that durations for the semi-automated methods (WS1, WS2) would grow at a slower pace than the manual method (WM). Additionally, the semi-automated method 2 (WS2) has potential to reduce duration more than the semi-automated method 1 (WS1) since there is only one comparison (WS2: a single discovered workflow of a group vs. the benchmark workflow) as opposed to multiple comparisons (WS1: individually discovered workflows of workflow instances vs. the benchmark workflow). Even though the discovered workflow for the semi-automated method 2 (WS2) may be more complex than WS1, conformance measurement duration was shorter since the repeated non-conformances were counted at once.

In Figure 4-4 and Figure 4-7, one of the most important modules that is optional in workflow conformance measurement is the performance measurement module (Module 2). In the aforementioned validation experiment, KPIs were not defined; thus, the performance measurement step was not fully addressed. For further discussion, the performance was assumed.

Success is defined as every activity meeting its own deadline (if there was an indicated deadline). That is, success was defined as every activity completing the task before the deadline (response_by). Thus, the dates and times recorded in the "completed_date_time" were supposed to be earlier than the corresponding dates and times of "response_by." When there was even one activity that was completed after the deadline or that was not completed at all (completed timestamp=NULL), the case was considered a failed performance.

For conformance measurement, the original benchmark workflow (Figure 6-9) was modified partially in order to create some conformant cases (since there were only non-conforming cases in Section 6.2.2). By adding parallel (|||) or sequential (≡) signs, some non-conformances were allowed in these new benchmark workflows (Benchmark Workflow revision 1 & 2 (Figure 6-16, magnified in Appendix G)). The parallel sign (|||) indicates that multiple participants may execute the activity concurrently. The sequential sign (≡) indicates that the activity may be completed consecutively.

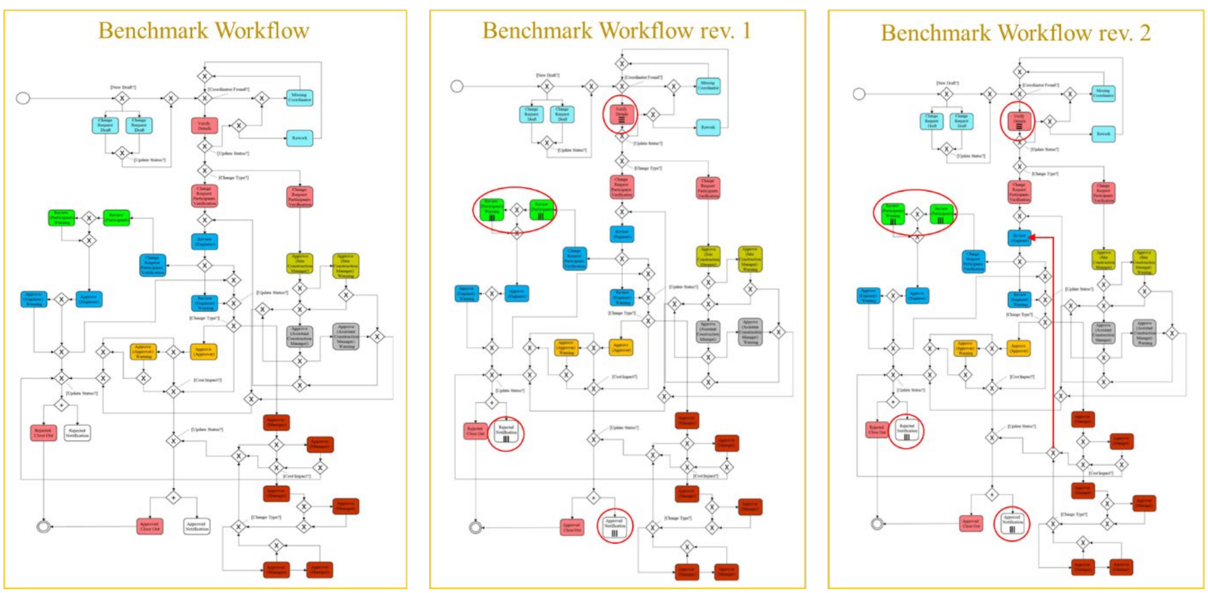


Figure 6-16: Examples of Benchmark Workflows (Original Benchmark Workflow, Revision 1 & 2)

There are four possible outcomes from measuring performance and conformance of a workflow instance. Their corresponding recommendations are described in Table 6-16. The X mark represents the outcome.

Table 6-16: Four Possible Outcomes and Corresponding Recommendations

Outcome Type	Success (S)	Failure (F)	Conformance (C)	Non-Conformance (N)	Recommendation
Outcome 1	X	-	X	-	Encourage the discovered workflow; remain the current benchmark workflow
Outcome 2	X	-	-	X	Consider modifying the benchmark workflow based on the discovered workflow
Outcome 3	-	X	X	-	Discourage the discovered workflow; consider modifying the benchmark workflow
Outcome 4	-	X	-	X	Remain the benchmark workflow until there is clearer relationship between conformance and performance

A discovered workflow is a result of process execution. In fact, the discovered workflow may not include every possible path of a benchmark workflow. Thus, the discovered workflow is either a part of the benchmark workflow (conformance) or a deviation of the benchmark workflow (non-conformance). When the discovered workflow is in conformance and the case closes successfully, it is recommended to encourage the discovered workflow and remain the current benchmark workflow (Outcome 1). When the discovered workflow is in conformance, but the case is not successful, it is recommended to discourage the discovered workflow and to consider modifying the current benchmark workflow (Outcome 3). When the discovered workflow is not in conformance and the case is successful, it is recommended to consider modifying the current benchmark workflow adapting the discovered workflow (Outcome 2). Lastly, when the discovered workflow is not in conformance and the case is not successful, it is recommended not to jump to any conclusion, but remain the benchmark workflow for further analysis. These recommendations can be provided when the performance measurement module (Module 2) is fully operated.

For validation, the semi-automated method 1 (WS1) approach was taken. For this experiment, when there was one flow path that was not in conformance, the entire case was considered not to be in conformance. The conformances of the eight cases were measured against either the benchmark workflow revision 1 or 2. Two benchmark workflows are illustrated in Appendix G. The difference between the benchmark workflow revision 1 and 2 is indicated with a red arrow. The corresponding benchmark workflow for each workflow instance was stated in the event log. Additionally, the performance was measured based on the event log. Results of the performance measurement module (Module 2) and the conformance measurement module (Module 3) are summarized in Table 6-17.

Table 6-17: The Results of Performance and Conformance Measurement (8 Workflow Instances)

Benchmark Workflow	Workflow Instance ID	Performance (S/F)	Conformance (C/N)	Outcome Type
Benchmark Workflow Revision 1	26	S	C	Outcome 1
	27	F	N	Outcome 4
	28	F	N	Outcome 4
	29	F	N	Outcome 4
Benchmark Workflow Revision 2	35	S	C	Outcome 1
	36	S	C	Outcome 1
	42	F	N	Outcome 4
	44	S	C	Outcome 1

The results implied that there may be a correlation between performance and conformance because success-conformance (S-C) and failure-non-conformance (F-N) were always matched together. However, further investigation is needed to make this conclusion. Two pieces of advice are given in this case. The first was to remain the successful discovered workflow within the benchmark workflow. This was derived from S-C outcome (Outcome 1). Because conformance to these paths (Workflow Instance 26, 35, 36, and 44) yielded successful performance, these paths were perceived adequate. An interesting finding was Workflow Instance 42. Workflow Instance 42 was relatively complex than other cases. It also utilized new path (marked as a red arrow from "Manager" to "Review (Engineer)" in Appendix G). This new path contributed to the conformance of Workflow Instance 42, though Workflow Instance 42 was concluded as a non-conformance (since any non-conformance led to non-conformance of a case).

The next piece of advice was to remain the benchmark workflow (Outcome 4). Although the performance failed, it was unknown whether the path in the benchmark workflow was suitable or not (due to its non-conformance). In this case, an auditor should not recommend to modify the benchmark workflow, but rather to remain it to achieve conformance. In this way, the performance of the benchmark workflow will be revealed in the future. Figure 6-17 presents where the non-conformances existed (marked with an "X"). When every path was in conformance, it was marked with a "V".

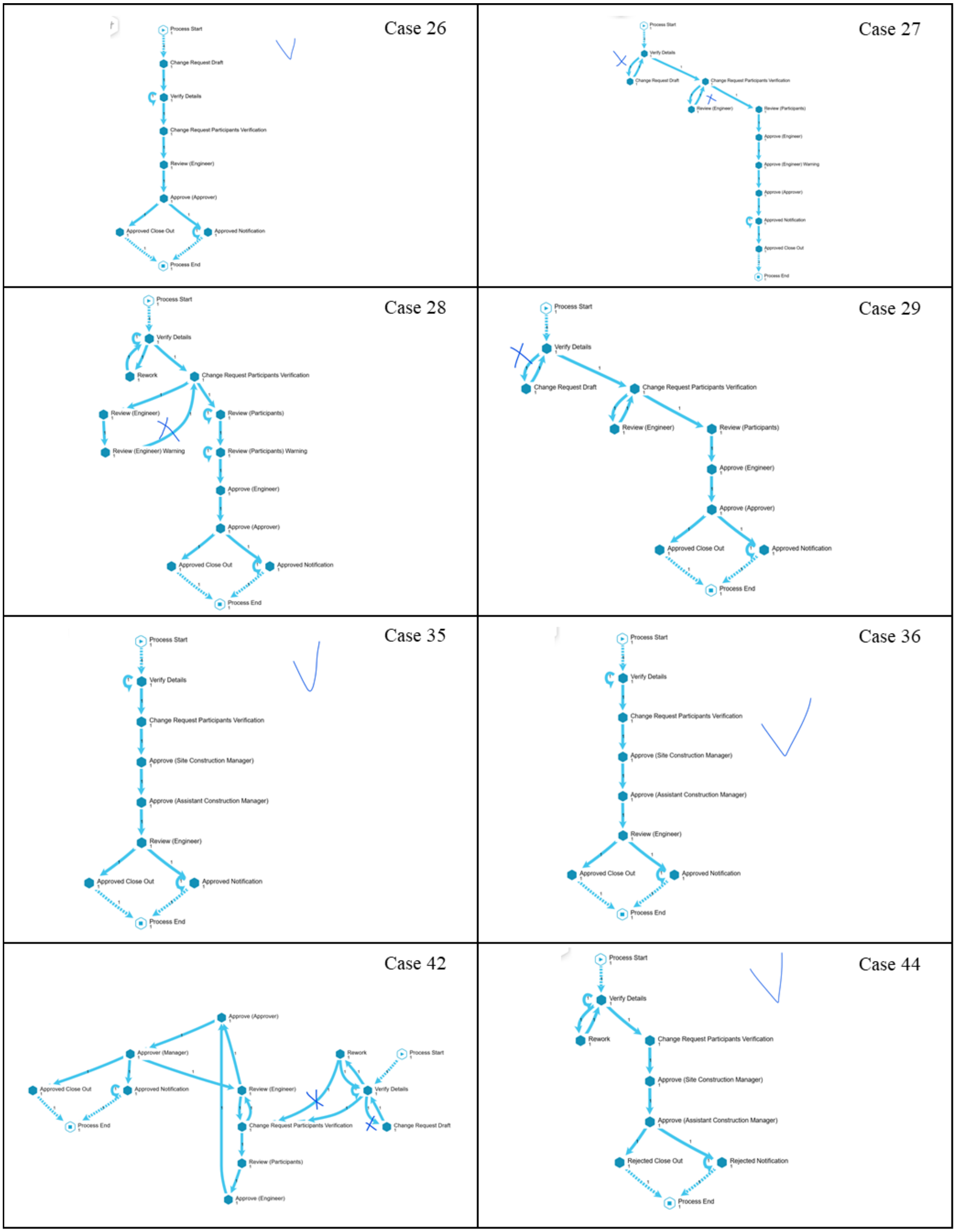


Figure 6-17: Existence of Non-Conformances by Case

Figure 6-17 clearly identifies where non-conformances occurred. Workflow Instances 26, 35, 36, and 44 were in conformance. Since parallel (|||) and sequential (≡) activities were allowed, these cases conformed with the benchmark workflows. Note that two types of benchmark workflows 1 and 2 were used for measuring non-conformances (Appendix G). Cases that were not in conformance were due to flow paths that were not allowed according to the benchmark workflows. For instance, "Verify Details" cannot be followed by "Change Request Draft."

This semi-experiment validated the need for the performance measurement module (Module 2) along with the main experiment which is presented in 6.2.2. The semi-automated workflow conformance measurement methods not only allowed efficient detection of non-conformances but also was able to provide reasonable advice on current process (i.e., workflow implementation).

Another issue is related to the definition of conformance since depending on the definition of conformance (Table 6-18), conformance rate will change. The conformance rate changes based on the defined benchmark workflow, and what is considered acceptable, for example, the non-existence of notification events. Revising the benchmark workflows allowed for exploring sensitivities of the conformance definitions. The revision changes are shown in red circles and arrows. In revision 1, parallel and sequential events were allowed. In revision 2, the conformance rate changed by allowing one flow path from "approver (manager)" to "review (engineer)."

Table 6-18: Conformance Definition Examples

Definition	Conformance Rate	Non-Conformance	NULL Timestamps	"Notification"	Benchmark Workflow	Method
1	Conformance Rate (CR) = Conforming workflow instances / Total workflow instances	Returns error	Not Replaced nor Removed	Non-existence Not allowed	Benchmark Workflow	WS1
2		One or more flow paths exist in the *Discovered Workflow* that do not exist in *Benchmark Workflow*	Replaced	Non-existence Not allowed	Benchmark Workflow	WS1
3			Removed	Non-existence Not allowed	Benchmark Workflow	WS1
4			Removed	Non-existence Allowed	Benchmark Workflow	WS1
5			Removed	Non-existence Allowed	Benchmark Workflow 2	WS1
6	CR = All correct flow paths / all flow paths		Removed	Non-existence Allowed	Benchmark Workflow	WS2

The highlights of Table 6-18 represent the sensitivity of the conformance rate definitions. All definitions inherit the rules and their changes from the previous definition.

- Definition 1: A workflow instance is non-conforming when the software returns an error due to the NULL timestamps of a workflow instance. If there is no error, the workflow instance is in conformance.
- Definition 2: NULL timestamps are replaced and the discovered workflow is compared against the original Benchmark Workflow.
- If one or more flow paths in the discovered workflow do not exist in the Benchmark Workflow, this workflow instance is a non-conforming workflow instance. Otherwise, the workflow instance is in conformance.
- Definition 3: Events of the NULL timestamps are removed rather than replaced.
- Definition 4: Non-existence of "Notification" is allowed.
- Definition 5: Benchmark Workflow revision 2 is used rather than the original Benchmark Workflow.

- Definition 6: the definition of conformance rate is different and semi-automated method 2 is used with the original benchmark workflow.

After automated process discovery and manual conformance checking, conformance rate is derived (Table 6-19).

Table 6-19: Results of the Conformance Rates (CR) based on Definition 1 - 6

Workflow Implemen -tation	**Workflow Instances**	**CR (%) = Conforming Workflow Instances / Total Workflow Instances**					**CR (%) = Conforming Flow Paths / Total Flow Paths**	
		Def. 1	**Def. 2**	**Def. 3**	**Def. 4**	**Def. 5**	**Def. 6**	
1	4	0	0	0	50.00	50.00	17/20	85.00
2	30	0	0	0	43.33	46.67	29/36	80.56
3	120	0	0	0	47.50	47.50	37/59	62.71
4	60	0	0	0	33.33	33.33	28/41	68.29
5	19	0	0	0	26.32	26.32	25/36	69.44
6	174	0	0	0	25.29	25.29	37/53	69.81
7	160	0	0	0	28.13	28.13	33/50	66.00
8	31	0	0	0	32.26	32.26	26/36	72.22

With definition 1, all workflow instances were non-conforming. Process discovery software did not run. With definition 2 and 3, the software ran, but there was no conforming workflow instance. The fourth and fifth definitions were adjusted to explore the sensitivity of conformance definitions. Exceptions were added to increase the conformance rate. Because benchmark workflow revision 2 allowed one more path, a non-conforming workflow instance with definition 4 is in conformance with definition 5. More than half of the workflow instances were still non-conforming with these exceptions. However, the data shows that there are patterns in the workflow instances since there are workflow instances that completely conform to the definition 4 or 5. The workflow implementation 1 was based upon but not identical to benchmark workflow. The business first enforced the workflow implementation 1. The business then adjusted their future workflow implementations 2 - 8 allowing some freedom within the control logic. Therefore, the conformance rate fluctuates based on the deviations from the benchmark workflow. The business attempts different strategies to enhance their process by adding automated events. However, the results deviate from the definition causing lower conformance rates at first. Eventually, the changes in the implementations start to improve the conformance. As the definitions are redefined, as shown in definition 6, the results can vary significantly. Even though the conformance rate is higher, it does not follow the same fluctuation as previous definitions. Two takeaways from Table 6-19 are that the semi-automated method 1 and 2 can be used to accomplish the computation of the conformance rate, and that auditors can analyze the business process depending on the definition of conformance.

Chapter 7 Conclusions and Future Work

This study aimed to semi-automate the practice conformance measurement process to assist an operational audit by suggesting a practice conformance model, a semi-automated framework, and toolsets. The objectives of discovering, testing, and evaluating the methods were achieved.

As a part of the toolset discovery, three software packages reached adequate accuracy and were selected to create a toolset and validate the framework. The semi-automated framework proposed by this study focused on two components from the conformance model (document and workflow). The document component with semi-automated methods produced up to 200 times quicker results not compromising accuracy. Companies can utilize the method and integrate their existing specialty software with customization. The workflow component with semi-automated methods reduced tedious work and increased efficiency. The conformance measurement allows the business to be proactive rather than reactive.

The semi-automated framework focused on document and workflow practice components from the practice conformance model. These two components were selected based on the literature and the commercial toolsets that were currently available. Commercial toolsets to semi-automate the measurement of practice conformance were found and examined in terms of adequate accuracy, ease of use, cost, and acceptability by project participants.

Some of the practice conformance elements such as "descriptions/narratives" and "process flowchart" were selected to achieve semi-automation measurements with commercial toolsets. The practice conformance model, semi-automated framework, and the toolsets jointly functioned as semi-automated practice conformance measurement methods.

For document conformance measurement, the industry, institute, and government documents were compared, and the methods that had different levels of automation were tested by identifying the usefulness of each module. Level 0 referred to the manual method. Level 1, Level 2, and Level 3 methods were semi-automated.

For workflow conformance measurement, it became clear that the semi-automated method can be used in the current situation to detect conformances quicker. There were possibilities of human errors for all methods including the manual method, but when the cases increase and the complexities also increase, it was perceived the semi-automated method 2 (WS2) was the quickest option. In the discussion, the semi-automated method 1 (WS1) was used to analyze individual cases to validate the need for the performance measurement module (Module 2). Module 2 was able to provide suggestions on how to improve the benchmark workflow.

The framework suggested by this study does not replace an in-person audit process. It is intended to assist the audit before the audit takes place. The semi-automated framework devised by this study is to increase the accuracy of the audit while reducing the time and cost associated with it. Thus, to be useful for the audit, the semi-automated framework must be streamlined. To streamline the framework, two things must be considered. First, modules included in the framework must have a solid purpose. Any module that is included in the framework takes time. Unless the benefits outweigh the costs, optional modules must be selected with caution. Second, integration among software packages must be considered. The transition time between modules increases when there are multiple modules. The integration of the toolsets proposed by this study will streamline the audit process significantly.

The framework suggested by this study successfully fulfilled the knowledge gap addressed in Chapter 2 by acquiring data from third-parties (documents) and machines (event logs). It also achieved semi-automation with the toolsets. The framework will be more robust with future work incorporated.

7.1 Conclusions

The results of this study have led to the key following conclusions:

1. Among fourteen commercialized toolsets that were evaluated, two document conformance measurement toolsets (Toolset 4 (WPS), Toolset 9 (Copyleaks)) and a workflow conformance measurement toolset (Toolset 14 (Celonis)) returned results that were accurate enough to assist auditors. These toolsets were able to achieve productivity and efficiency by reducing time and human errors.

2. The semi-automated methods presented by this study reduced duration of document conformance measurement from 3 to 200 times compared to the manual method. Levels of automation affected the results of accuracy, costs, and time. In general, reducing the manual work and increasing the automation level saved time and costs. Since the manual work does not always increase accuracy, it is recommended to measure the practice conformance using the method with highest level of automation since the higher level of automation takes lesser time.

3. The semi-automated methods presented by this study reduced duration of workflow conformance measurement by more than half compared to the manual method. It was concluded that functions such as process discovery from the process mining software reduced a significant amount of tedious work, thus, increasing efficiency.

By automating some tedious preparation tasks before the actual analysis, an auditor can focus more on important issues (typically those that require higher cognitive skills such as judgment) and allocate time more productively. For example, comparing stacks of documents and complex workflows are essential but inefficient and error-prone when completed manually. Samples for an audit have often been collected based on an auditor's knowledge or experience. However, the auditor's judgment may overlook some potential problems that need to be addressed. Due to the deadlines and lack of resources, everything cannot be manually taken into consideration.

A semi-automated framework covers the entire body, not only samples. For example, such a framework considers an entire body of texts when it comes to the keyword extraction or text matching module. The framework considers every case when it comes to comparing an event log to a benchmark workflow. Thus, the semi-automated framework identifies the areas where an auditor should be investigating by detecting anomalies or exceptions. This allows for selecting appropriate samples, saving time, and making the best use of resources. In summary, a semi-automated practice conformance measurement framework can increase the efficiency of the audit thereby potentially mitigating audit fees between auditors and audited companies. This study validated that with the current states of technology, practice conformance measurement can be semi-automated.

7.2 Contributions

The main contributions of this thesis can be summarized in four aspects:

1. Confirmation of the utility of semi-automated practice conformance measurement, as a part of an audit

A semi-automated conformance measurement suggested by this study is a useful step prior to the manual assessment of auditors. While there have been difficulties in comparing data among construction companies, this study is more feasible as it only requires internal data (e.g., event logs and benchmark workflows). This study contributes to the body of knowledge and to the practices for companies needing to analyze large amounts of data in a short period with reasonable costs and accuracy.

Full automation is not a reality since the technology is not available for sale. Semi-automation, on the other hand, can be reached and can assist human audits more realistically. Automating tedious and repetitive tasks is similar to the chauffeur (assistant) model approach, and preventing big mistakes is similar to the guardian model approach in the automotive industry. Like the two semi-automation models, this thesis provided solutions to reduce tedious and repetitive human tasks and to find human errors that were missed by the manual method.

2. Establishment of a simple practice conformance model

A practice conformance model of a capital project was established based on a review of the literature and current practice. To analyze the practice conformance, data was collected based on the grounded theory research methodology. The practice conformance model was finalized with an iterative process between logical inductive analyses and systematically obtained data analysis.

In Figure 7-1, beneath each of the four components, the left checklist represents how to measure practice conformance and the right bullets represent conformance elements. Existence of the conformance elements, to some degree, provides evidence of conformance. In this study, document and workflow components were perceived to be semi-automated, but, because all four components are interrelated, this study can be a steppingstone for people and action component approach as well.

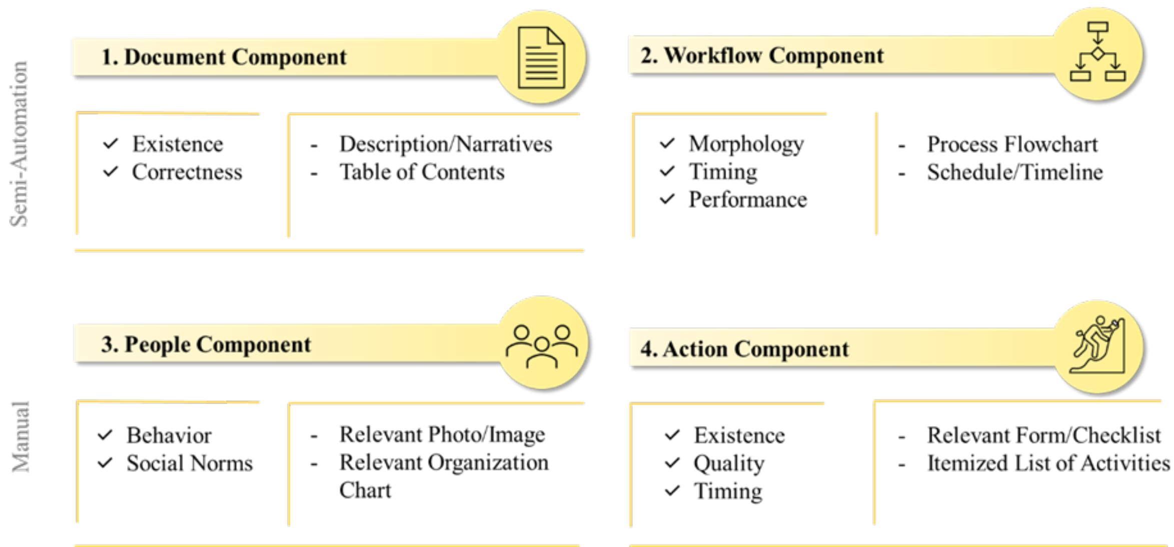


Figure 7-1: Practice Conformance Model Establishment

3. Development and validation of a consistent and repeatable practice conformance measurement framework and identification of areas that can be (semi-) automated

One of the challenges in previous works was that a single set of metrics was not suitable for all construction industry sectors to use. Surveys could assist as they were often cheaper and easier to understand, but the surveys can be subjective when self-reported. Especially, since BM&M programs usually took a long time to build, there were limitations to accommodate all the demands from the industry. It took time not only to develop such programs but also to introduce them to potential participants, get their feedback, revise the tools, re-introduce them, and get the commitment from firms to provide data on an ongoing basis (Caldas et al., 2015; Nasir et al., 2012). They also typically required industry champions which this study did not require. Therefore, along with robust conformance model, the strength of this research is the conformance measurement framework (Figure 7-2). Two methods were developed from this framework, document being the first method and workflow being the second. Moreover, the framework suggested by this study allows customization depending on the needs and definition of conformance. The two conformance components (document and workflow) were measured with a structured framework and methods that allowed semi-automation (in yellow) with the toolsets that reached adequate accuracy.

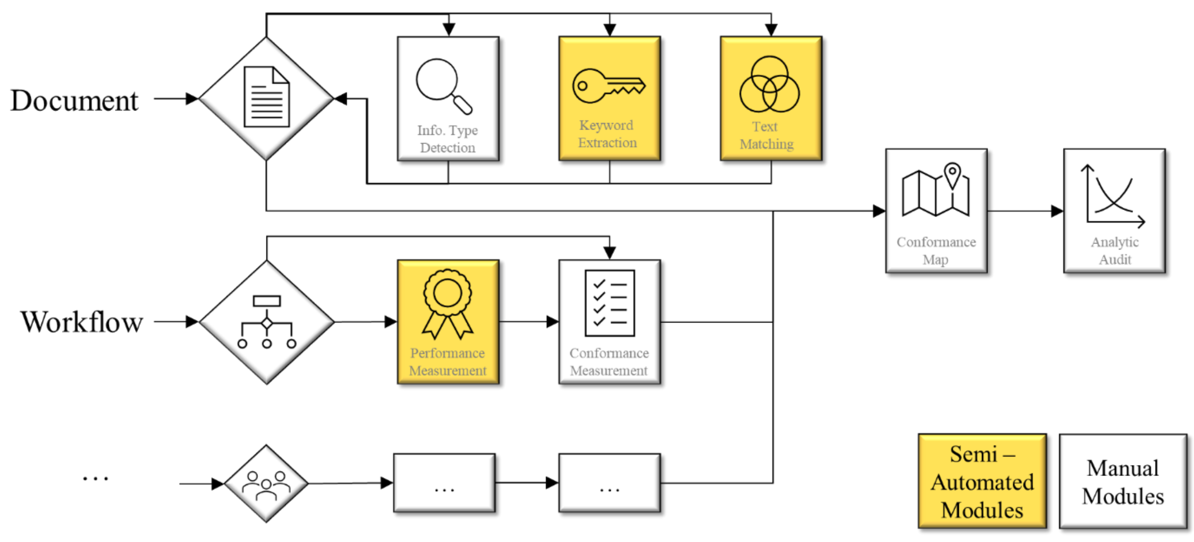


Figure 7-2: Practice Conformance Measurement Framework

The document conformance measurement method allows the auditor to select modules and customize the order based on their own experience or judgment. This method is preparation for the analytic audit so that non-conformances are detected to determine where to focus on.

The workflow conformance measurement method involves process discovery function included in Module 2 and conformance checking function included in Module 3 to measure conformance. The workflow conformance measurement method can be used for sampling or detecting non-conformance types.

4. Discovery and application of commercialized toolsets that are accurate enough to be utilized for the semi-automated practice conformance measurement

In previous research studies, obtaining data from contractors was perceived as a challenge. This was due to the lack of communication with the subcontractors, confidentiality of data, and no control regarding the provision of data from subcontractors. Also, the contractor community was hesitant to provide information on productivity and performance measures due to concerns regarding confidentiality and a conviction in many cases that they have nothing to learn. This study, on the other hand, does not require big data or the development of machine learning algorithms. Though software packages may have embedded machine learning algorithms (e.g., text matching), the auditor does not have to build the artificial intelligence (AI) training model to compute the data. Commercial software packages are already available to use for this purpose.

7.3 Limitations

The objectives, the model, framework, and toolsets proposed with this work have reached expectations. The capital project practice conformance model provided the practice conformance components and measurable practice conformance elements that became the foundation of the framework. The toolsets were able to save time and produce repeatable results. However, this study has limitations, which can be grouped in four categories:

1. The toolsets suggested in this study have technical limitations and areas that can still be improved.

When it comes to converting PDF or image files, texts were not translated accurately when the font was not embedded in the applications. When a document was handwritten, translation to the system was not as accurate as PDF conversion. For the workflow conformance measurement, many of the non-conformances that were detected by the software were inaccurate.

2. Not all practice conformance elements defined by this study were tested.

For instance, the "schedule/timeline" element had the highest potential impact when automated (Section 3.2.2.2). There are many software packages for automatic scheduling. However, conformance measurement for the "schedule/timeline" element was not tested in this thesis.

3. The document and workflow components inherently have weaknesses.

The document component has disadvantages since documents can be manipulated easily. As an alternative, workflow conformance has been suggested to be measured. Similar to documents, however, event log data, which is the raw data of workflow conformance measurement, also has inherent weaknesses as not all human behaviours or actions appear in the event log. The event log also can be manipulated or misinterpreted because a few clicks can create an event. For example, if a participant sends a document accidentally to another participant that is not expecting the document, the event log will have unexpected long duration since the recipient never took an action on the document. It is not accurate because the task required less time than recorded. More limitations of this research exist in terms of scope and methodology, since time and resources of the candidate were constrained. Some of those limitations are indirectly addressed in the following section on recommendations for future work.

4. Unclear distinction between human and machine control

In some cases, the distinction between human- and machine-control is unclear in this thesis. For example, hand-off problems of transitioning from machine-control to human-control are challenging. However, this thesis focused on human and machine collaboration (semi-automation), by defining each module and handing off the conformance maps as an outcome before analytic audit.

7.4 Recommendations for Future Work

The following recommendations are suggested based on this thesis:

- Integration with document management system (DMS) and workflow management system (WMS)

State-of-the-art information systems typically include a document management system (DMS) and a workflow management system (WMS). The document conformance measurement framework suggested by this study can also be a part of a DMS, because they share common goals and functionalities, such as checking the existence and correctness of elements. A DMS is an electronic system to store, track, and manage documents for organization and accessibility. For example, with a DMS, if files need to be attached, the system notifies the users to attach a file when they are missing. If the framework and toolsets suggested in this thesis are embedded in a DMS, the system will be able to check whether the correct file has been uploaded.

Additionally, a DMS provides multiple features, such as searching, monitoring, versioning, indexing, data validating, and publishing. These features are also interrelated to document conformance measurement and can be further developed along with this study. For example, versioning tools from DMS track edits to documents and recover older versions. The toolsets suggested can further develop the versioning tools by adding functions such as text matching.

A DMS commonly has either a built-in workflow module or close integration with the workflow management system (WMS). The Workflow Management System (WMS) is an electronic system that automatically routes the data to the predetermined tasks. Whereas the manual workflow requires users to decide whom to send the data to, an automated workflow or a WMS automatically sends the data to pre-set participants. The WMS can be integrated with process mining to measure the efficiency of the workflow or to improve the workflow.

- Incorporation of the state-of-the-art equipment into practice conformance measurement

Critical literature review and a variety of consultations have assisted in developing the practice conformance model presented in this thesis. However, the practice conformance model does not incorporate equipment, such as BIM, laser scanners, drones, virtual reality, and augmented reality technologies. These technologies are used to build digital twins or BIM models and to measure geometric conformances between the as-is model and the to-be model. Although geometric conformance is not considered as practice conformance in this study, the technologies from the literature review can be used to extend the scope of the conformance model in the future.

- Semi-automation of "people" and "actions" components of the practice conformance model

Though the event log from the workflow component account for parts of human behaviours or actions, not all human behaviours or actions are reflected in the event log. Therefore, people and actions components remain as future work. To measure the conformance of people and action components, practice conformance element such as a "relevant project form/checklist" that is created in automated ways can be used. For instance, a drone can detect workers who are not wearing personal protective equipment and record the non-conformance on a "relevant project form/checklist" with image mining technology in automated ways.

- Semi-automated conformance measurement of other practice conformance elements

The semi-automated practice conformance framework requires software packages. In this thesis, among thirteen practice conformance elements, only two elements (i.e., "description/narratives," "process flowchart") were elaborated and tested. The other eleven practice conformance elements (e.g., "schedule/timeline") also have high potential to be semi-automated (Section 3.2.2.3). By utilizing technologies such as text mining, process mining, and image mining, more practice conformance elements may be measured. Well-trained artificial intelligence (AI) systems that are embedded in the software packages will lead to higher accuracy. Thus, each element can be a topic for further research (e.g., image mining for “relevant photos/images” element).

- Increased automation in manual modules

Currently, manual modules exist in the semi-automation framework. For example, information type detection module is completed manually. The module requires detecting thirteen elements and then grouping them accordingly. For example, the accuracy of machines distinguishing an image of “document approval/authorization” practice conformance element and an image of “relevant photos/images” practice conformance element is not high. Thus, technology should be developed further to increase accuracy.

- Improvement in document and workflow conformance maps

There are opportunities to improve both document and workflow conformance maps. As an analogy, a geographical map has a function to zoom in and out to guide more accurately based on the user’s needs. With this functionality in mind, there can be more general or specific conformance maps in the future. In a document conformance map, overall or specific parts of conformance can be expressed. For instance, when zoomed in, matching phrases are highlighted while when zoomed out, matching paragraphs are highlighted.

References

Ackoff, R. L. (1989). From Data to Wisdom. *Journal of Applied System Analysis*, *16*(1), 3–9.

Acur, N., Gertsen, F., Sun, H., & Frick, J. (2003). The formalisation of manufacturing strategy and its influence on the relationship between competitive objectives, improvement goals, and action plans. *International Journal of Operations & Production Management*, *23*(10), 1114–1141. https://doi.org/10.1108/01443570310496599

Adriansyah, A., Van Dongen, B. F., & Van der Aalst, W. M. P. (2011). Conformance Checking Using Cost-Based Fitness Analysis. *2011 IEEE 15th International Enterprise Distributed Object Computing Conference*, 55–64. https://doi.org/10.1109/EDOC.2011.12

Ahn, S., Shokri, S., Lee, S., Haas, C. T., & Haas, R. C. G. (2017). Exploratory Study on the Effectiveness of Interface-Management Practices in Dealing with Project Complexity in Large-Scale Engineering and Construction Projects. *Journal of Management in Engineering*, *33*(2), 04016039. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000488

Armas-Cervantes, A., Baldan, P., Dumas, M., & Garcia-Bañuelos, L. (2016). Diagnosing behavioral differences between business process models: An approach based on event structures. *Information Systems*, *56*, 304–325. https://doi.org/10.1016/j.is.2015.09.009

Ashford, J. L. (2002). *The Management of Quality in Construction*. Routledge.

Bingham, E. (2010). *Development of the project definition rating index (PDRI) for infrastructure projects*. Arizona State University.

Blaser, R., Schnabel, M., Biber, C., Bäumlein, M., Heger, O., Beyer, M., Opitz, E., Lenz, R., & Kuhn, K. A. (2007). Improving pathway compliance and clinician performance by using information technology. *International Journal of Medical Informatics*, *76*(2), 151–156. https://doi.org/10.1016/j.ijmedinf.2006.07.006

Buchgeher, G., & Weinreich, R. (2013). Towards Continuous Reference Architecture Conformance Analysis. In K. Drira (Ed.), *Software Architecture* (pp. 332–335). Springer. https://doi.org/10.1007/978-3-642-39031-9_32

Buijs, J. C. A. M., Van Dongen, B. F., & Van der Aalst, W. M. P. (2012). On the Role of Fitness, Precision, Generalization and Simplicity in Process Discovery. In R. Meersman, H. Panetto, T. Dillon, S. Rinderle-Ma, P. Dadam, X. Zhou, S. Pearson, A. Ferscha, S. Bergamaschi, & I. F. Cruz (Eds.), *On the Move to Meaningful Internet Systems: OTM 2012* (pp. 305–322). Springer. https://doi.org/10.1007/978-3-642-33606-5_19

Caldas, C. H., Kim, J., Haas, C. T., Goodrum, P. M., & Zhang, D. (2015). Method to Assess the Level of Implementation of Productivity Practices on Industrial Projects. *Journal of Construction Engineering and Management*, *141*(1), 04014061. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000919

Carnes, W. R. (1984). *Signature verification system* (United States Patent No. US4433436A). https://patents.google.com/patent/US4433436A/en

Cha, H., & O'Connor, J. T. (2005). Optimizing Implementation of Value Management Processes for Capital Projects. *Journal of Construction Engineering and Management*, *131*(2), 239–251. https://doi.org/10.1061/(ASCE)0733-9364(2005)131:2(239)

Chalechale, A., Naghdy, G., Premaratne, P., & Mertins, A. (2004). Document image analysis and verification using cursive signature. *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)*, *2*, 887–890. https://doi.org/10.1109/ICME.2004.1394343

Chang, A. P., Chou, C. C., Lin, J. D., & Hsu, C. Y. (2013). Road Construction Project Environmental Impact Assessment Scope Definition Using Project Definition Rating Index (PDRI). *Advanced Materials Research*, *723*, 885–892. https://doi.org/10.4028/www.scientific.net/AMR.723.885

Chen, L., Kang, S., Karimidorabati, S., & Haas, C. T. (2019). Improving the Quality of Event Logs in the Construction Industry for Process Mining. *2019 Proceedings of the 36th ISARC*, *36*, 804–811. https://doi.org/10.22260/ISARC2019/0108

Cho, C., & Gibson, Jr, G. E. (2000). Development of a Project Definition Rating Index (PDRI) for General Building Projects. *Construction Congress VI: Building Together for a Better Tomorrow in an Increasingly Complex World*, 343–352. https://doi.org/10.1061/40475(278)38

Collot, R., Achemlal, M., & Revillet, M. (1991). *Signature verification method and system with optimization of static parameters* (United States Patent No. US5042073A). https://patents.google.com/patent/US5042073A/en

Deng, P. S., Liao, H. M., Ho, C. W., & Tyan, H. (1999). Wavelet-Based Off-Line Handwritten Signature Verification. *Computer Vision and Image Understanding*, *76*(3), 173–190. https://doi.org/10.1006/cviu.1999.0799

Deshayes, L., Foufou, S., & Gruninger, M. (2007). An Ontology Architecture for Standards Integration and Conformance in Manufacturing. In S. Tichkiewitch, M. Tollenaere, & P. Ray (Eds.), *Advances in Integrated Design and Manufacturing in Mechanical Engineering II* (pp. 261–276). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6761-7_18

Duffy, K., Ferguson, C., & Watson, H. (2004). Data collecting in grounded theory-some practical issues. *Nurse Researcher*, *11*(4).

Eastman, C., Lee, J., Jeong, Y., & Lee, J. (2009). Automatic rule-based checking of building designs. *Automation in Construction*, *18*(8), 1011–1033. https://doi.org/10.1016/j.autcon.2009.07.002

García-Bañuelos, L., Van Beest, N. R. T. P., Dumas, M., Rosa, M. L., & Mertens, W. (2018). Complete and Interpretable Conformance Checking of Business Processes. *IEEE Transactions on Software Engineering*, *44*(3), 262–290. https://doi.org/10.1109/TSE.2017.2668418

Gideon, S. J., Kandulna, A., Kujur, A. A., Diana, A., & Raimond, K. (2018). Handwritten Signature Forgery Detection using Convolutional Neural Networks. *Procedia Computer Science*, *143*, 978–987. https://doi.org/10.1016/j.procs.2018.10.336

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Adline de Gruyter.

Golzarpoor, B. (2017). *Industry Foundation Processes (IFP): Theoretical and Practical Foundations for the Construction Industry*. University of Waterloo.

Golzarpoor, B., Haas, C. T., & Rayside, D. (2016). Improving process conformance with Industry Foundation Processes (IFP). *Advanced Engineering Informatics*, *30*(2), 143–156. https://doi.org/10.1016/j.aei.2016.02.005

Gündüz Murat, Nielsen Yasemin, & Özdemir Mustafa. (2013). Quantification of Delay Factors Using the Relative Importance Index Method for Construction Projects in Turkey. *Journal of Management in Engineering*, *29*(2), 133–139. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000129

Hussein, W., Salama, M. A., & Ibrahim, O. (2016). Image Processing Based Signature Verification Technique to Reduce Fraud in Financial Institutions. *MATEC Web of Conferences*, *76*, 05004. https://doi.org/10.1051/matecconf/20167605004

Hwang, B., Thomas, S. R., Haas, C. T., & Caldas, C. H. (2009). Measuring the Impact of Rework on Construction Cost Performance. *Journal of Construction Engineering and Management*, *135*(3), 187–198. https://doi.org/10.1061/(ASCE)0733-9364(2009)135:3(187)

International Organization for Standardization. (2015). Quality management systems — Fundamentals and vocabulary (ISO Standard No: 9000:2015). https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:en

Karimidorabati, S. (2014). *A Model for Implementing an Automated Change Management Process for Construction Mega-projects.* University of Waterloo.

Kang, S., & Haas, C. T. (2018). Evaluating Artificial Intelligence Tools for Automated Practice Conformance Checking. *2018 Proceedings of the 35th ISARC*, *35*, 110–117. https://doi.org/10.22260/ISARC2018/0015

Kaur, N., Pushe, V., & Kaur, R. (2014). Natural Language Processing Interface for Synonym. *International Journal of Computer Science and Mobile Computing*, *3*(7), 638–642.

KPMG (2020) 20 key risks to consider by Internal Audit before 2020

Laugen, B. T., Acur, N., Boer, H., & Frick, J. (2005). Best manufacturing practices: What do the best-performing companies do? *Emerald Group Publishing Limited*, *25*(2), 131–150. https://doi.org/info:doi/10.1108/01443570510577001

Lucas, J., Bulbul, T., & Anumba, C. (2013). Gap Analysis on the Ability of Guidelines and Standards to Support the Performance of Healthcare Facilities. *Journal of Performance of Constructed Facilities*, *27*(6), 748–755. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000364

Maani, K. E., & Sluti, D. G. (1990). A Conformance — Performance Model: Linking Quality Strategies to Business Unit's Performance. In J. E. Ettlie, M. C. Burstein, & A. Fiegenbaum (Eds.), *Manufacturing Strategy: The Research Agenda for the Next Decade Proceedings of the Joint industry University Conference on Manufacturing Strategy Held in Ann Arbor, Michigan on January 8–9, 1990* (pp. 85–96). Springer Netherlands. https://doi.org/10.1007/978-94-009-2189-4_10

Mannhardt, F. (2018). *Multi-perspective Process Mining*. Eindhoven University of Technology.

Mariscal, G., Marbán, Ó., & Fernández, C. (2010). A survey of data mining and knowledge discovery process models and methodologies. *The Knowledge Engineering Review*, *25*(2), 137–166. https://doi.org/10.1017/S0269888910000032

Meng, X. (2012). The effect of relationship management on project performance in construction. *International Journal of Project Management*, *30*(2), 188–198. https://doi.org/10.1016/j.ijproman.2011.04.002

Mettyear, N. (2016). *Dynamic handwriting verification and handwriting-based user authentication* (United States Patent No. US9235748B2). https://patents.google.com/patent/US9235748B2/en

Mosadeghrad, A. M. (2013). Healthcare service quality: Towards a broad definition. *International Journal of Health Care Quality Assurance*, *26*(3), 203–219. https://doi.org/10.1108/09526861311311409

Nasir, H., Haas, C. T., Rankin, J. H., Fayek, A. R., Forgues, D., & Ruwanpura, J. (2012). Development and implementation of a benchmarking and metrics program for construction performance and productivity improvement. *Canadian Journal of Civil Engineering*, *39*(9), 957–967. https://doi.org/10.1139/l2012-030

Newman, K., & Cowling, A. (1996). Service quality in retail banking: The experience of two British clearing banks. *International Journal of Bank Marketing*, *14*(6), 3–11. https://doi.org/10.1108/02652329610130127

NHTSA (2020) https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety

Nielsen, J., & Landauer, T. K. (1993, May). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* 206-213.

O'Malley, A. S., Clancy, C., Thompson, J., Korabathina, R., & Meyer, G. S. (2004). Clinical Practice Guidelines and Performance Indicators as Related—But Often Misunderstood—Tools. *The Joint Commission Journal on Quality and Safety*, *30*, 48–56. https://doi.org/10.1016/S1549-3741(04)30109-7

Patlakas, P., Livingstone, A., Hairstans, R., & Neighbour, G. (2018). Automatic code compliance with multi-dimensional data fitting in a BIM context. *Advanced Engineering Informatics*, *38*, 216–231. https://doi.org/10.1016/j.aei.2018.07.002

Pelland, D. (2017). *2017 Audit Fee Survey Report*. Financial Executive Research Foundation (FERF).

Rahman, M. M. (2014). Barriers of Implementing Modern Methods of Construction. *Journal of Management in Engineering*, *30*(1), 69–77. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000173

Rankin, J., Fayek, A. R., Meade, G., Haas, C. T., & Manseau, A. (2008). Initial metrics and pilot program results for measuring the performance of the Canadian construction industry. *Canadian Journal of Civil Engineering*, *35*(9), 894–907. https://doi.org/10.1139/L08-018

Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2016). Process mining in healthcare: A literature review. *Journal of Biomedical Informatics*, *61*, 224–236. https://doi.org/10.1016/j.jbi.2016.04.007

Roth, A. V., & Jackson, W. E. (1995). Strategic Determinants of Service Quality and Performance: Evidence from the Banking Industry. *Management Science*, *41*(11), 1720–1733. https://doi.org/10.1287/mnsc.41.11.1720

Rozinat, A. (2005). *Conformance Testing: Measuring the Alignment Between Event Logs and Process Models*. Eindhoven University of Technology.

Rozinat, A., & Van der Aalst, W. M. P. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, *33*(1), 64–95. https://doi.org/10.1016/j.is.2007.07.001

Rudyi, S. S., Vovk, T. A., & Rozhdestvensky, Y. V. (2019). Signature identification by Minkowski dimension. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *29*(5), 053110. https://doi.org/10.1063/1.5092270

Salama, D. A., & El-Gohary, N. M. (2013). Automated Compliance Checking of Construction Operation Plans Using a Deontology for the Construction Domain. *Journal of Computing in Civil Engineering*, *27*(6), 681–698. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000298

Salama, D. M., & El-Gohary, N. M. (2016). Semantic Text Classification for Supporting Automated Compliance Checking in Construction. *Journal of Computing in Civil Engineering*, *30*(1), 04014106. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000301

Salama, D. M., & El-Gohary, N. M. (2011). Semantic Modeling for Automated Compliance Checking. *Computing in Civil Engineering (2011)*, 641–648. https://doi.org/10.1061/41182(416)79

Shan, Y., Goodrum, P. M., Zhai, D., Haas, C. T., & Caldas, C. H. (2011). The impact of management practices on mechanical construction productivity. *Construction Management and Economics*, *29*(3), 305–316. https://doi.org/10.1080/01446193.2010.538070

Siddiqi, S., & Sharan, A. (2015). Keyword and keyphrase extraction techniques: a literature review. *International Journal of Computer Applications*, *109*(2).

Suriadi, S., Andrews, R., ter Hofstede, A. H. M., & Wynn, M. T. (2017). Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs. *Information Systems*, *64*, 132–150. https://doi.org/10.1016/j.is.2016.07.011

Thomas, H. R., Mathews, C. T., & Ward, J. G. (1986). Learning Curve Models of Construction Productivity. *Journal of Construction Engineering and Management*, *112*(2), 245–258. https://doi.org/10.1061/(ASCE)0733-9364(1986)112:2(245)

Van Beest, N. R. T. P., Dumas, M., García-Bañuelos, L., & La Rosa, M. (2015). Log Delta Analysis: Interpretable Differencing of Business Process Event Logs. In H. R. Motahari-Nezhad, J. Recker, & M. Weidlich (Eds.), *Business Process Management* (pp. 386–405). Springer International Publishing. https://doi.org/10.1007/978-3-319-23063-4_26

Van Berlo, L. A., & Natrop, M. (2015). BIM on the construction site: Providing hidden information on task specific drawings. *Journal of Information Technology in Construction (ITcon)*, *20*(7), 97–106.

Van der Aalst, W. M. P. (2017). Spreadsheets for business process management: Using process mining to deal with "events" rather than "numbers"? *Business Process Management Journal*, *24*(1), 105–127. https://doi.org/10.1108/BPMJ-10-2016-0190

Van der Aalst, W. M. P., Van Hee, K. M., Van der Werf, J. M., & Verdonk, M. (2010). Auditing 2.0: Using Process Mining to Support Tomorrow's Auditor. *IEEE*, *43*(3), 90–93. https://doi.org/10.1109/MC.2010.61

Van Schaijk, S., & Van Berlo, L. A. H. M. (2016). Introducing process mining for AECFM: Three experimental case studies. *Proceedings of the 11th European Conference on Product and Process Modelling (ECPPM 2016), Limassol, Cyprus*, 481–486.

Wang, Y. (2002). *Applying the PDRI in project risk management.* The University of Texas at Austin

Weijters, A. J. M. M., & Van der Aalst, W. M. P. (2001). Process mining: Discovering workflow models from event-based data. *Belgium-Netherlands Conf. on Artificial Intelligence*.

Werner, M. (2017). Financial process mining—Accounting data structure dependent control flow inference. *International Journal of Accounting Information Systems*, *25*, 57–80. https://doi.org/10.1016/j.accinf.2017.03.004

Wilson, D. L., Baddeley, A. J., & Owens, R. A. (1997). A New Metric for Grey-Scale Image Comparison. *International Journal of Computer Vision*, *24*(1), 5–17.

Yun, S., Choi, J., De Oliveira, D. P., & Mulva, S. P. (2016). Development of performance metrics for phase-based capital project benchmarking. *International Journal of Project Management*, *34*(3), 389–402. https://doi.org/10.1016/j.ijproman.2015.12.004

Yurchyshyna, A., Faron-Zucker, C., Le Thanh, N., & Zarli, A. (2010). Knowledge capitalisation and organisation for conformance checking model in construction. *International Journal of Knowledge Engineering and Soft Data Paradigms*, *2*(1), 15–32.

Yurchyshyna, A., & Zarli, A. (2009). An ontology-based approach for formalisation and semantic organisation of conformance requirements in construction. *Automation in Construction*, *18*(8), 1084–1098. https://doi.org/10.1016/j.autcon.2009.07.008

Zhai, D., Goodrum, P. M., Haas, C. T., & Caldas, C. H. (2009). Relationship between Automation and Integration of Construction Information Systems and Labor Productivity. *Journal of Construction Engineering and Management*, *135*(8), 746–753. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000024

Zhang, J., & El-Gohary, N. (2012). Extraction of Construction Regulatory Requirements from Textual Documents Using Natural Language Processing Techniques. *Computing in Civil Engineering (2012)*, 453–460. https://doi.org/10.1061/9780784412343.0057

Zhang, J., & El-Gohary, N. M. (2015). Automated Information Transformation for Automated Regulatory Compliance Checking in Construction. *Journal of Computing in Civil Engineering*, *29*(4), B4015001. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000427

Zhang J., & El-Gohary, N. M. (2016). Semantic NLP-Based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking. *Journal of Computing in Civil Engineering*, *30*(2), 04015014. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000346

Zhang, J., & El-Gohary, N. M. (2017a). Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. *Automation in Construction*, *73*, 45–57. https://doi.org/10.1016/j.autcon.2016.08.027

Zhang, J., & El-Gohary, N. M. (2017b). Semantic-Based Logic Representation and Reasoning for Automated Regulatory Compliance Checking. *Journal of Computing in Civil Engineering*, *31*(1), 04016037. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000583

Zhang, J., & El-Gohary, N. M. (2013). Information Transformation and Automated Reasoning for Automated Compliance Checking in Construction. *Computing in Civil Engineering (2013)*, 701–708. https://doi.org/10.1061/9780784413029.088

Zhang, L., & Ashuri, B. (2018). BIM log mining: Discovering social networks. *Automation in Construction*, *91*, 31–43. https://doi.org/10.1016/j.autcon.2018.03.009

Zhong, B., Gan, C., Luo, H., & Xing, X. (2018). Ontology-based framework for building environmental monitoring and compliance checking under BIM environment. *Building and Environment*, *141*, 127–142. https://doi.org/10.1016/j.buildenv.2018.05.046

Zhou, P., & El-Gohary, N. M. (2017). Ontology-based automated information extraction from building energy conservation codes. *Automation in Construction*, *74*, 103–117. https://doi.org/10.1016/j.autcon.2016.09.004

Appendix A: Experimental Results of Text Extraction (Hand-Written Document Conversion Software Packages)

Four toolsets that have the potential to convert hand-written documents to machine-readable texts are tested and results are shown below. Microsoft Word does not have this function, but Google OCR, Adobe Pro,WPS have the functionality. Google OCR has the best result among others; however, it is still not accurate enough to use it for commercial use.

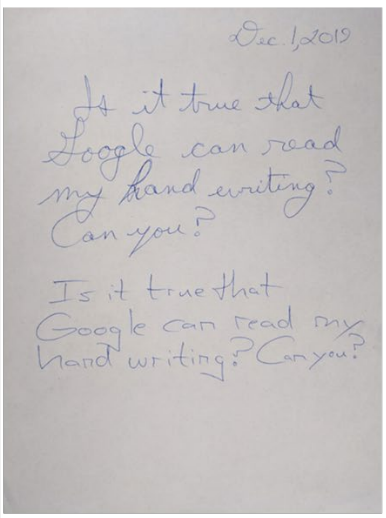
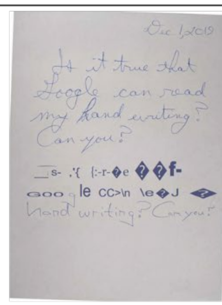
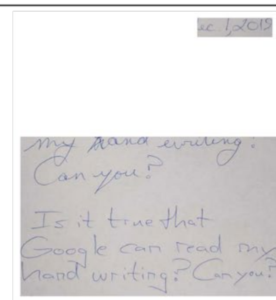
<table>
<tr><th>Original Image</th><th>Toolset 1
(Google OCR)</th><th>Toolset 2
(Microsoft Word)</th></tr>
<tr><td rowspan="3">
</td><td>Dec 1, 2010 Is it true that Google can read my hand ewritings
Can your Is it true that
Google can read my hand writing? Can you?</td><td></td></tr>
<tr><th>Toolset 3
(Adobe Pro)</th><th>Toolset 4 (WPS)</th></tr>
<tr><td></td><td></td></tr>
</table>

Figure A-1: Experimental Results of Software for Hand-Written Document Conversion

Appendix B: Graphical Results of Keyword Extraction (Word Cloud Generator Software Packages)

A document of “Interface Management” is used for creating a word cloud with four word cloud generator software packages. Results below imply that while some toolsets (e.g., WordArt) can stem and lemmatize partially, others (e.g., WordClouds) cannot. Because the keyword extraction toolsets remove stopwords and present words visually, hidden keywords may be revealed through this module.

<table>
<tr><th>Toolset 5 (WordClouds)</th><th>Toolset 6 (WordArt)</th></tr>
<tr><td>


</td><td>


</td></tr>
<tr><th>Toolset 7 (TagCrowd)</th><th>Toolset 8 (WordItOut)</th></tr>
<tr><td>


</td><td>


</td></tr>
</table>

Figure B-1: Experimental Results of Software for Keyword Extraction

Appendix C: Original Documents for Text Matching

Twenty-two documents are compared for the text matching. Below are the two excerpts of documents. “Social Situations” and “Consequences of a Policy Violation” sections are used for the functional demonstration in Chapter 5.

Document A

4.0 SOCIAL SITUATIONS

4.1 Company Sponsored Social Events

In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community. Responsible alcohol use is permitted at Company sponsored social functions held away from Company premises, which must have the prior approval of a Vice President level or above, and will be conducted in accordance with the Company's hosting guidelines. Anyone who attends and consumes alcohol must not be returning to or going to work after the event.

4.2 Business Hosting

Consistent with the above, if alcohol is made available to guests in the course of conducting business (e.g. restaurant meeting), employees are expected to use judgment and be responsible in hosting others.

5.0 CONSEQUENCES OF A POLICY VIOLATION

5.1 General Requirements

Any violation of the provisions of this Policy may result in corrective action or termination of employment. Management has the authority and discretion to hold out of service any individual who is believed to be involved in an incident that could lead to corrective action pending the results of the investigation. The appropriate action in a particular case depends on the nature of the policy violation and the circumstances surrounding the situation; the severity of the violation may warrant entering the corrective action process at different levels or termination of employment. For all employees any confirmed situation of drug trafficking on Company premises will result in termination of employment.

A positive drug test, or an alcohol test result of .04 BAC or higher, or a refusal to test are all considered a violation of this Policy. An alcohol test of .02 BAC, or higher, for anyone working at a designated dry site is considered a violation of this Policy. In all other situations, an employee who has an alcohol test result of .02 to .039 BAC in a reasonable cause or post incident situation will be removed from duty until considered safe to return (at a minimum not before their next work day or shift) and may be subject to corrective action.

After any confirmed violation, the employee may be referred for a SAE assessment to determine whether there is a need for a structured treatment program.

Document B

8. SOCIAL SITUATIONS

In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community. Subject to any site specific limitations, responsible Alcohol use may be permitted at Company sponsored social functions with appropriate prior approval. Alternative transportation arrangements will be made available when possible.

Consistent with the above, if Alcohol is made available to Company guests in the course of conducting Company Business (e.g., restaurant meetings), Employees are expected to use reasonable judgment and be responsible in hosting others, and remain in compliance with the Policy and Supporting Standards.

Procedures for hosting events are set out in the Social and Business Hosting Standard.

9. CONSEQUENCES OF A POLICY VIOLATION

a) General Requirements: Any violation of this Policy and Supporting Standards may result in discipline up to and including termination of employment. In all situations, an investigation will be conducted to verify that a Policy or Standard violation has occurred. The appropriate discipline in a particular case depends on the nature of the Policy or Standard violation and the circumstances surrounding the situation. The severity of the violation will warrant entering the discipline process at different levels. General violations of this Policy include:

i. failure to comply with the Policy and Supporting Standards;

ii. a positive Alcohol or Drug test (refer to the Alcohol and Drug Testing Standard); or

iii. a Failure to Test.

b) Referral for Assessment: After any confirmed positive Alcohol and Drug test, an Employee may be referred by Health and Wellness to a Substance Abuse Professional for a Substance Abuse Assessment (refer to the Substance Abuse Assessment Standard). Failing to meet with the Substance Abuse Professional or attend a scheduled Substance Abuse Assessment is a violation of this Policy.

c) Conditions for Continued Employment: Should the Company determine that employment will be continued after a violation of the Policy or Supporting Standards, the Employee will be required to enter into an agreement governing their continued employment which may require any or all of the following actions, or any other condition appropriate to the situation:

i. temporary removal from their position;

ii. adherence to any recommended treatment and aftercare program;

iii. successful completion of a return to work Alcohol and Drug test;

iv. ongoing unannounced follow-up Alcohol and Drug testing for the duration of their agreement;

v. adherence to any ongoing rehabilitation conditions or requirements; and

vi. no further Policy or Standard violations during the monitoring period.

Failure to meet the requirements of the agreement will be grounds for discipline up to and including termination.

Figure C-1: Excerpts of the Original Documents for Text Matching

Appendix D: Experimental Results of Text Matching (Other Software Packages)

The entire performance of Toolset 9 (Copyleaks) and Toolset 10 (Copyscape) are presented in Chapter 5. Additionally, a part of results from Toolset 11 (PlagScan) and Toolset 12 (CountWordsFree) are presented below. Toolset 11 (PlagScan) is intentded for crosschecking against the web; thus, it is inefficient for the purpose of this study where particular documents comparison is the specific needs. Toolset 12 (CountWordsFree) does not provide as accurate results as the other toolsets.

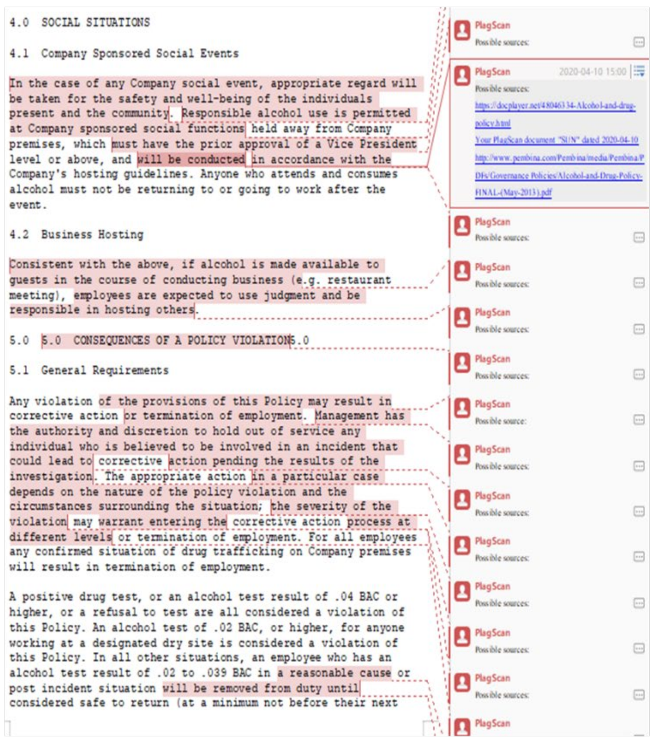
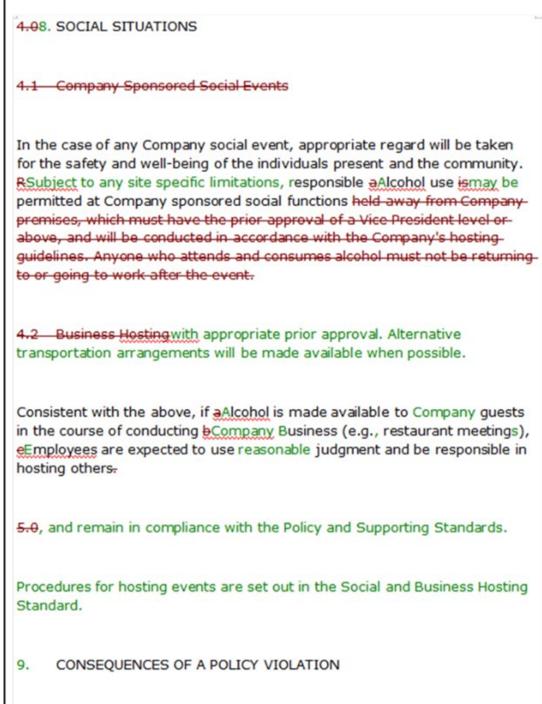
<table>
<tr><th>Toolset 11 (PlagScan)</th><th>Toolset 12 (CountWordsFree)</th></tr>
<tr><td>

4.0 SOCIAL SITUATIONS
4.1 Company Sponsored Social Events
In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community. Responsible alcohol use is permitted at Company sponsored social functions held away from Company premises, which must have the prior approval of a Vice President level or above, and will be conducted in accordance with the Company's hosting guidelines. Anyone who attends and consumes alcohol must not be returning to or going to work after the event.
4.2 Business Hosting
Consistent with the above, if alcohol is made available to guests in the course of conducting business (e.g. restaurant meeting), employees are expected to use judgment and be responsible in hosting others.
5.0 5.0 CONSEQUENCES OF A POLICY VIOLATION5.0
5.1 General Requirements
Any violation of the provisions of this Policy may result in corrective action or termination of employment. Management has the authority and discretion to hold out of service any individual who is believed to be involved in an incident that could lead to corrective action pending the results of the investigation. The appropriate action in a particular case depends on the nature of the policy violation and the circumstances surrounding the situation; the severity of the violation may warrant entering the corrective action process at different levels or termination of employment. For all employees any confirmed situation of drug trafficking on Company premises will result in termination of employment.
A positive drug test, or an alcohol test result of .04 BAC or higher, or a refusal to test are all considered a violation of this Policy. An alcohol test of .02 BAC, or higher, for anyone working at a designated dry site is considered a violation of this Policy. In all other situations, an employee who has an alcohol test result of .02 to .039 BAC in a reasonable cause or post incident situation will be removed from duty until considered safe to return (at a minimum not before their next
PlagScan
Possible sources:
PlagScan 2020-04-10 15:00
Possible sources:
https://docplayer.net/48046334-Alcohol-and-drug-policy.html
Your PlagScan document "SI/N" dated 2020-04-10
http://www.pembina.com/Pembina/media/Pembina/PDFs/Governance Policies/Alcohol-and-Drug-Policy-FINAL-(May-2013).pdf
PlagScan
Possible sources:
</td><td>

~~4.0~~8. SOCIAL SITUATIONS
~~4.1 Company Sponsored Social Events~~
In the case of any Company social event, appropriate regard will be taken for the safety and well-being of the individuals present and the community. ~~R~~Subject to any site specific limitations, responsible ~~a~~Alcohol use ~~is~~may be permitted at Company sponsored social functions ~~held away from Company premises, which must have the prior approval of a Vice President level or above, and will be conducted in accordance with the Company's hosting guidelines. Anyone who attends and consumes alcohol must not be returning to or going to work after the event.~~
~~4.2 Business Hosting~~with appropriate prior approval. Alternative transportation arrangements will be made available when possible.
Consistent with the above, if ~~a~~Alcohol is made available to Company guests in the course of conducting ~~b~~Company Business (e.g., restaurant meetings), ~~e~~Employees are expected to use reasonable judgment and be responsible in hosting others~~.~~
~~5.0~~, and remain in compliance with the Policy and Supporting Standards.
Procedures for hosting events are set out in the Social and Business Hosting Standard.
9. CONSEQUENCES OF A POLICY VIOLATION
</td></tr>
</table>

Figure D-1: Experimental Results of Text Matching

Appendix E: An Event Log Data Used for Functional Demonstration

In total, there are 598 workflow instnaces and 16633 activities. For functional demonstrations, two workflow instances are used. Below event log shows the two workflow instances. For Figure 5-1, workflow instance 27 (WF_ID: 27) is the example of an individual workflow instance. For Figure 5-2, workflow instances 26 and 27 are grouped which is used for the experiment.

Table E-1: An Event Log (Workflow Instances 26, 27)

WF_ID	Activity_Display_Name	Created_Date_Time	Completed_Date_Time
26	Change Request Draft	3/3/2011 0:34	3/10/2011 1:31
26	Verify Details	3/9/2011 18:31	3/10/2011 1:31
26	Verify Details	3/10/2011 1:31	3/10/2011 14:15
26	Change Request Participants Verification	3/10/2011 14:15	3/10/2011 14:58
26	Review (Engineer)	3/10/2011 14:58	3/10/2011 15:12
26	Approve (Approver)	3/10/2011 15:12	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Notification	3/10/2011 15:13	3/10/2011 15:13
26	Approved Close Out	3/10/2011 15:13	3/10/2011 15:27
27	Verify Details	3/3/2011 12:48	3/3/2011 16:44
27	Change Request Draft	3/3/2011 16:44	3/3/2011 19:48
27	Verify Details	3/3/2011 19:48	3/3/2011 20:29
27	Change Request Participants Verification	3/3/2011 20:29	3/3/2011 20:36
27	Review (Engineer)	3/3/2011 20:36	3/3/2011 21:00

27	Change Request Participants Verification	3/3/2011 21:00	3/3/2011 21:01
27	Review (Participants)	3/3/2011 21:01	3/3/2011 22:35
27	Review (Participants)	3/3/2011 21:01	3/4/2011 18:09
27	Review (Participants)	3/3/2011 21:01	3/3/2011 21:07
27	Review (Participants)	3/3/2011 21:01	3/3/2011 22:18
27	Review (Participants)	3/3/2011 21:01	3/8/2011 16:25
27	Review (Participants)	3/3/2011 21:01	3/3/2011 21:32
27	Review (Participants)	3/3/2011 21:01	3/3/2011 21:09
27	Approve (Engineer)	3/8/2011 19:16	3/9/2011 17:00
27	Approve (Engineer) Warning	3/9/2011 17:00	3/9/2011 18:47
27	Approve (Approver)	3/9/2011 18:47	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Notification	3/10/2011 15:16	3/10/2011 15:16
27	Approved Close Out	3/10/2011 15:16	3/10/2011 17:12

Appendix F: Eight Versions of Workflow Implementations of a Change Request Process

Below are the eight change request process workflows. The eight versions are different from each other slightly and the changes are circled in red. The event log reflects all eight versions of the workflow. Although most differences in the change request process workflows do not affect the BPMN, there is an exception in the second workflow. For example, due to the creation of "UpdateStatusInReview," "Approver (Manager)" may be followed by "Review (Engineer)."

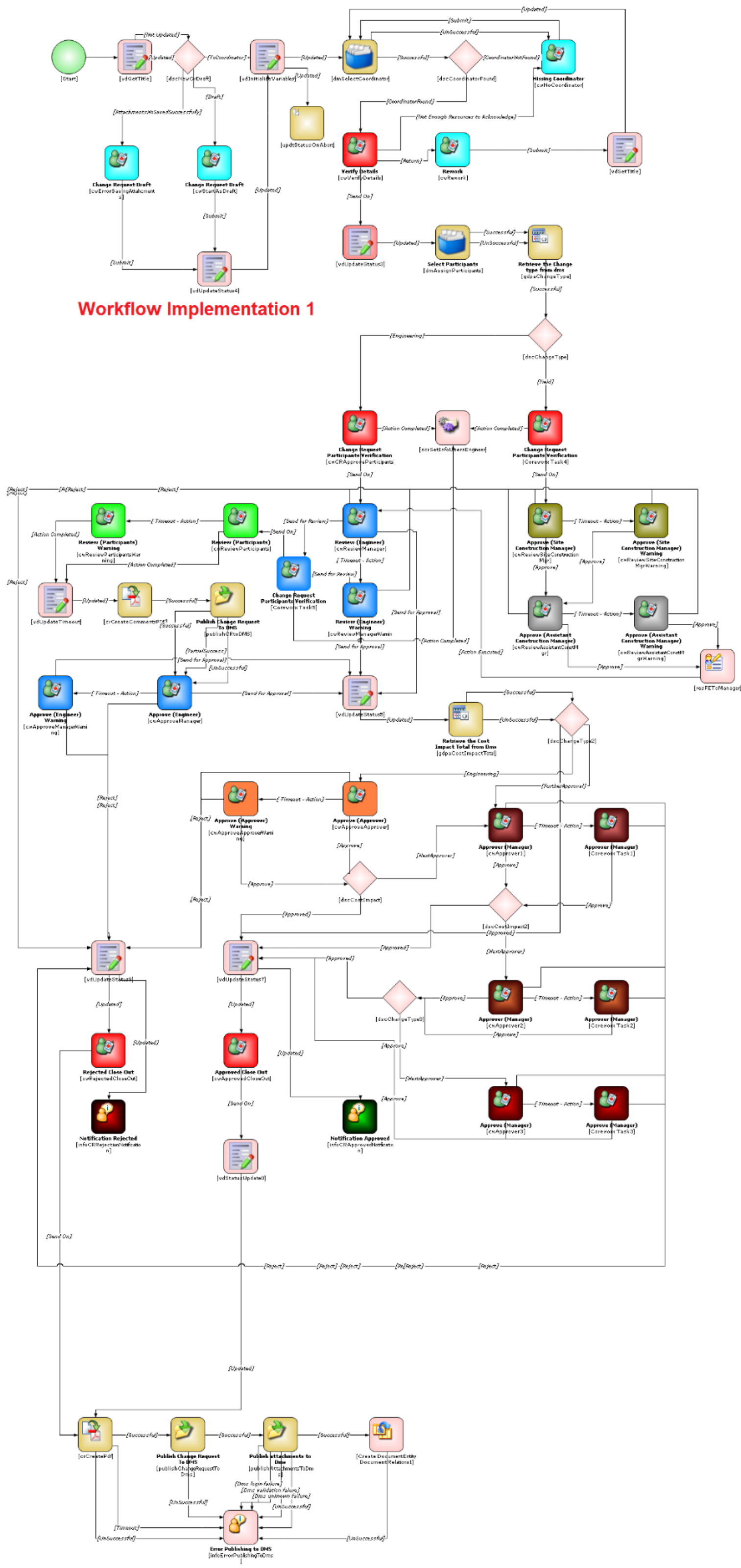


Figure F-1: Workflow Implementation 1

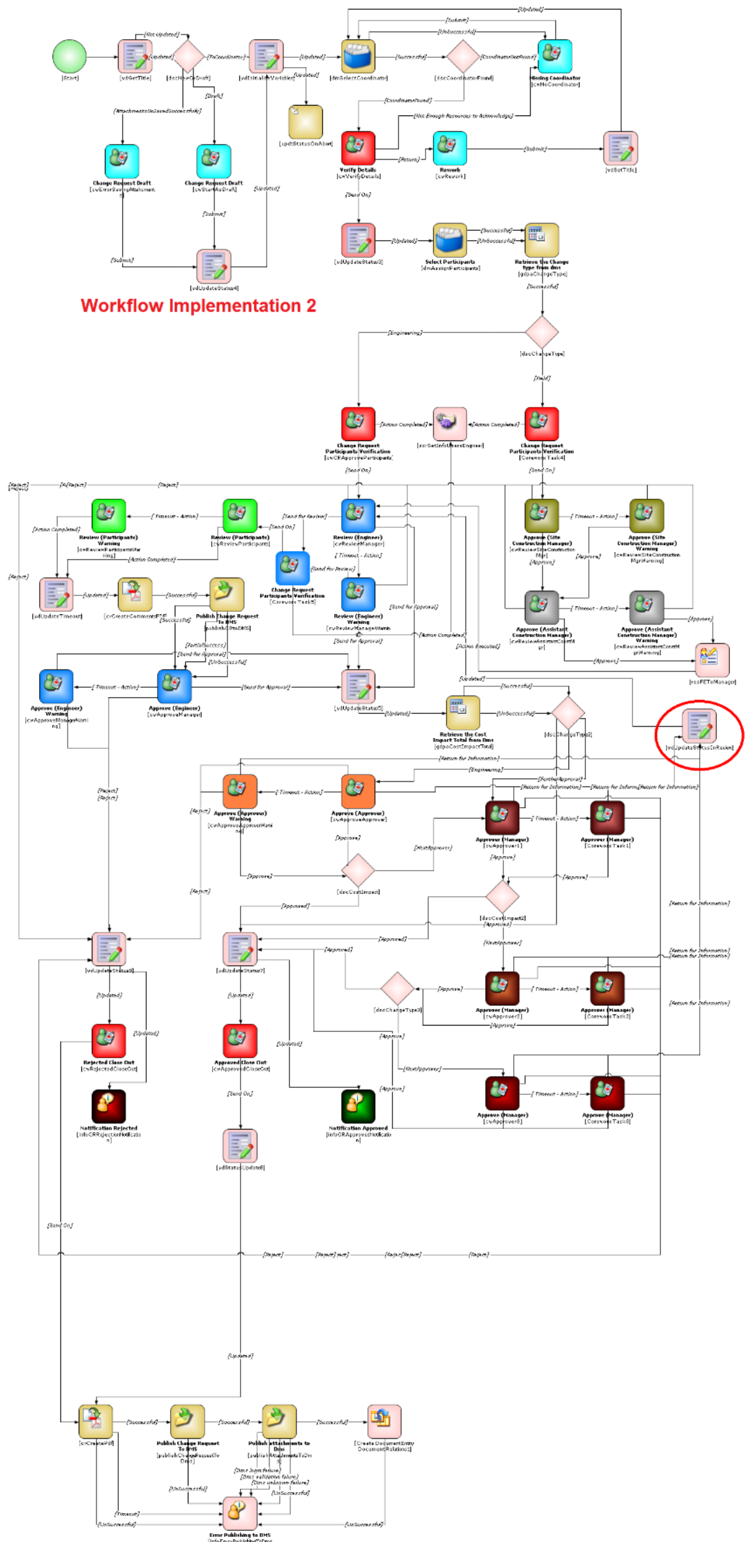


Figure F-2: Workflow Implementation 2

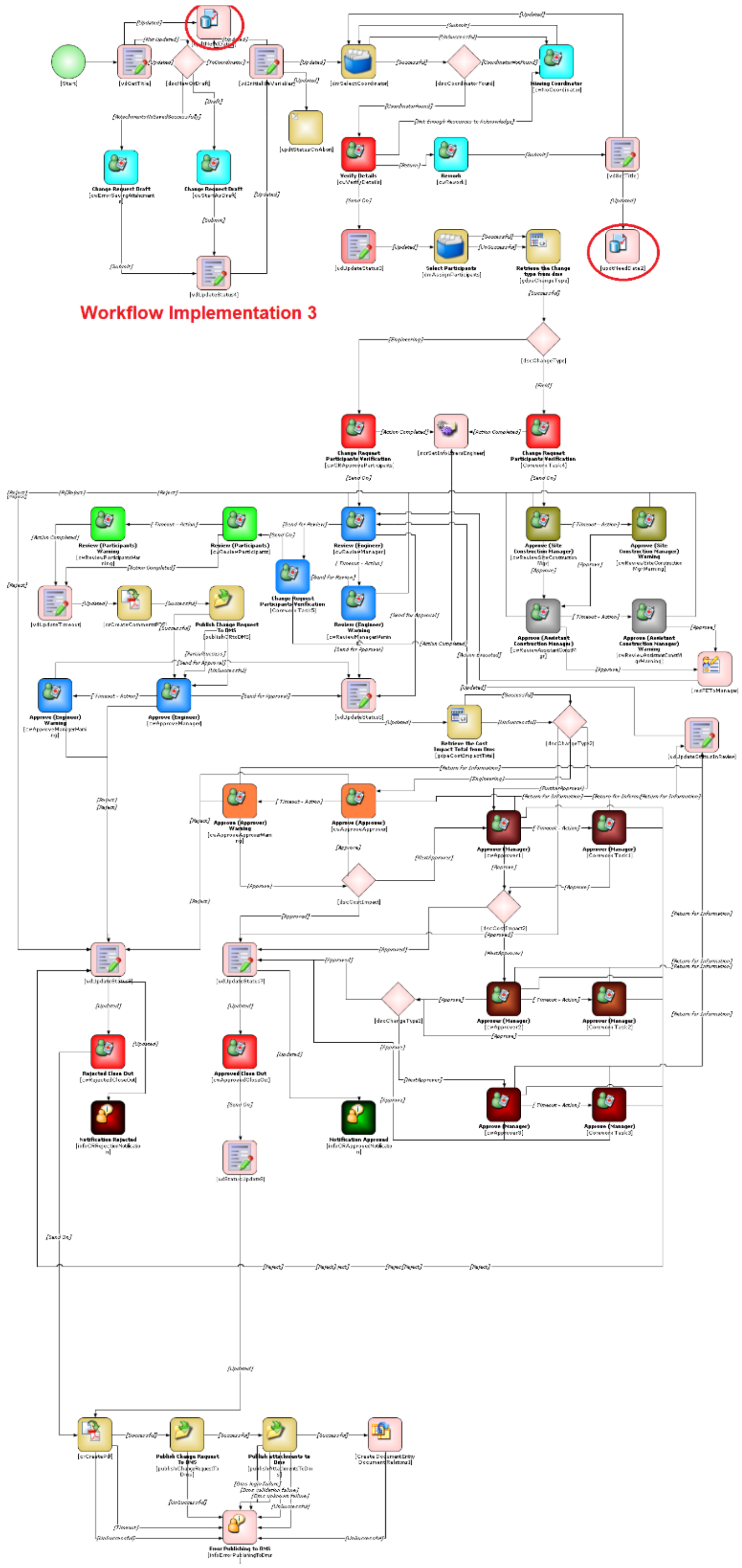


Figure F-3: Workflow Implementation 3

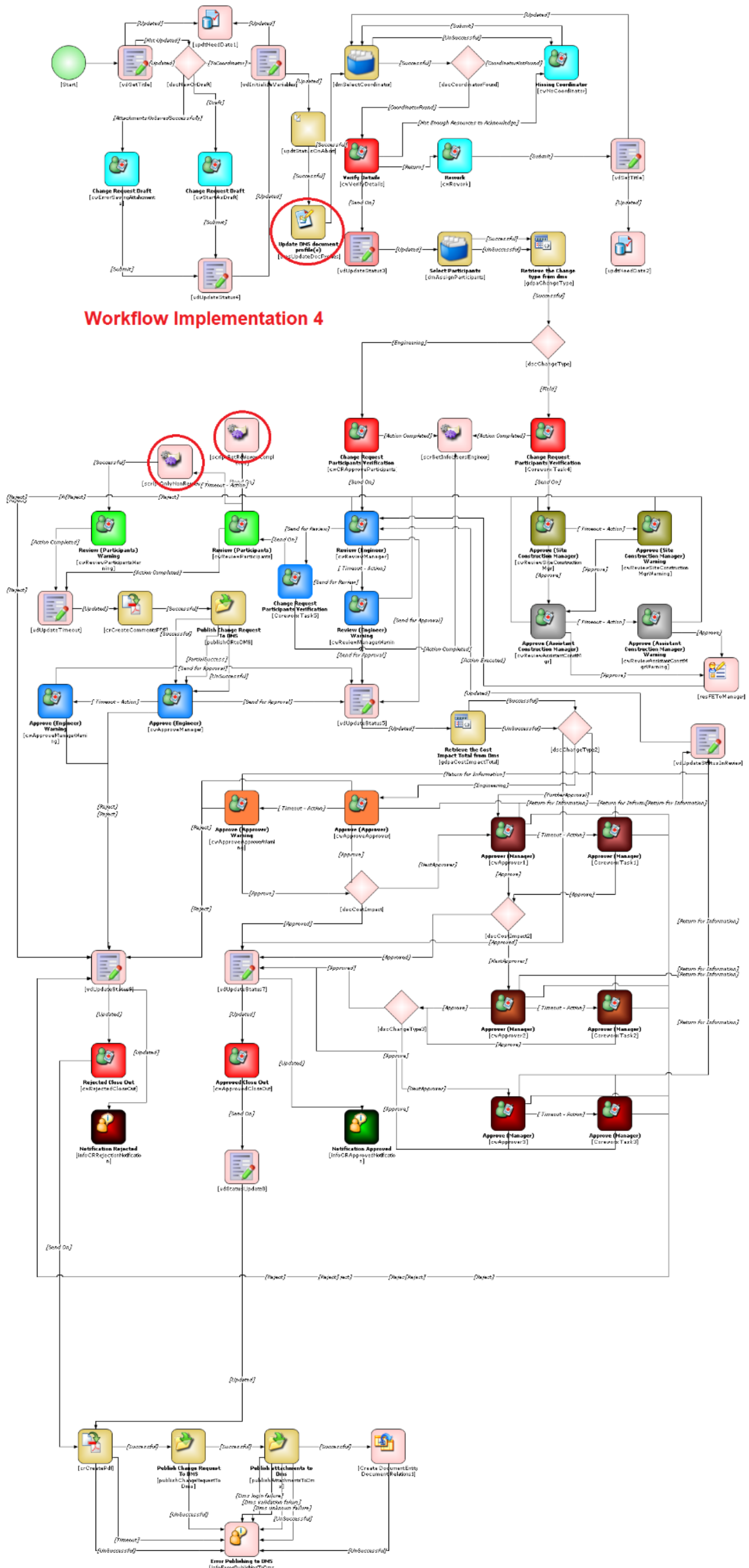


Figure F-4: Workflow Implementation 4

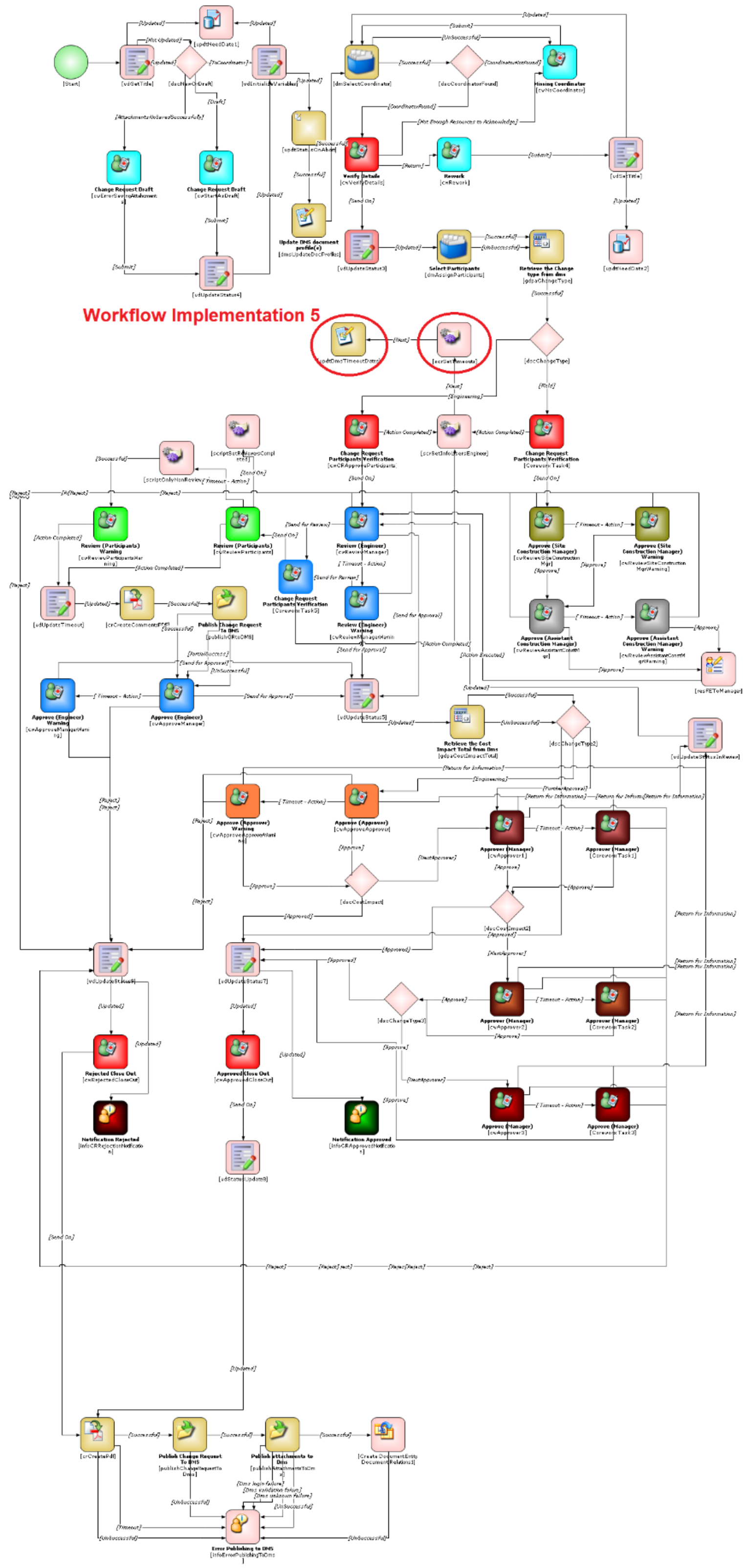


Figure F-5: Workflow Implementation 5

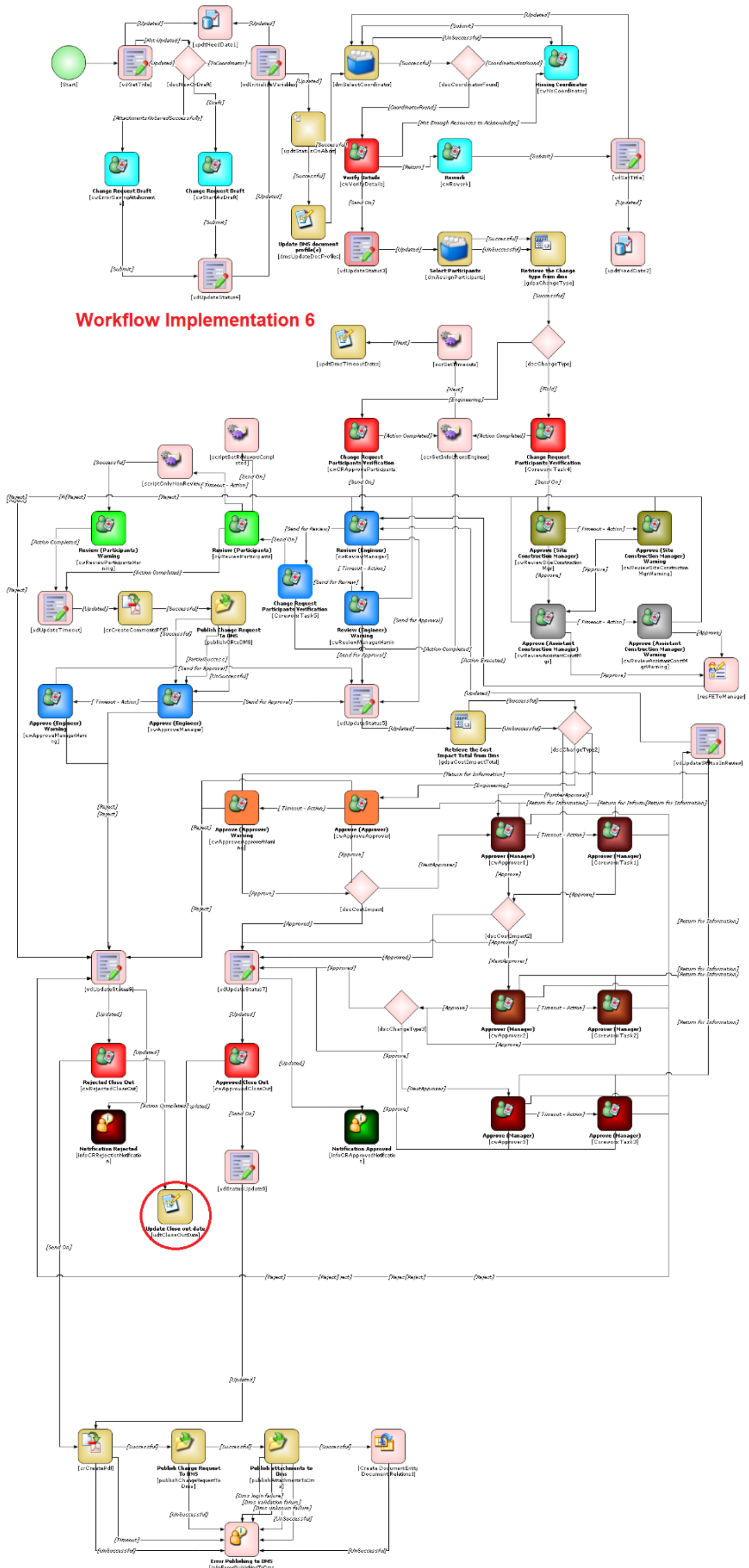


Figure F-6: Workflow Implementation 6

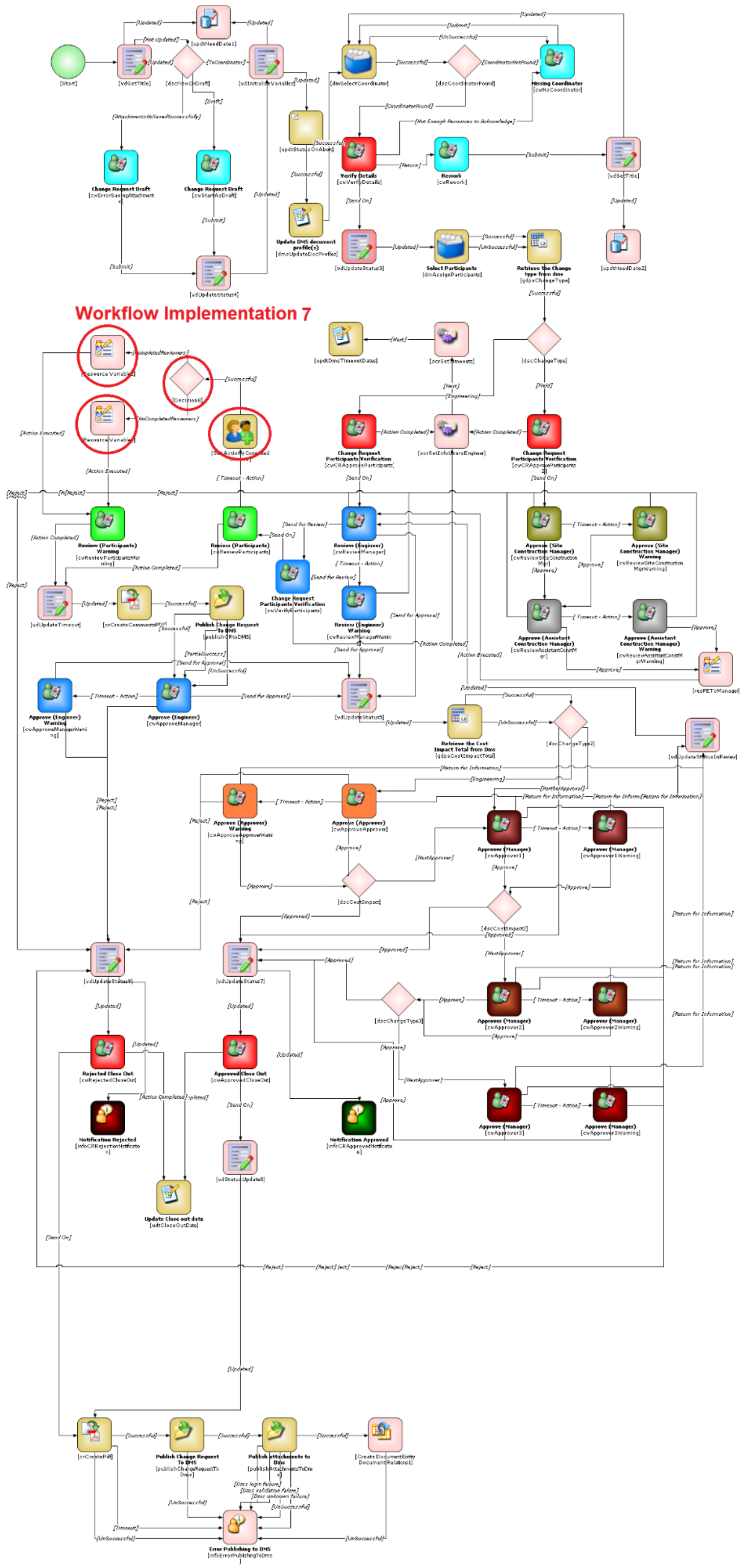


Figure F-7: Workflow Implementation 7

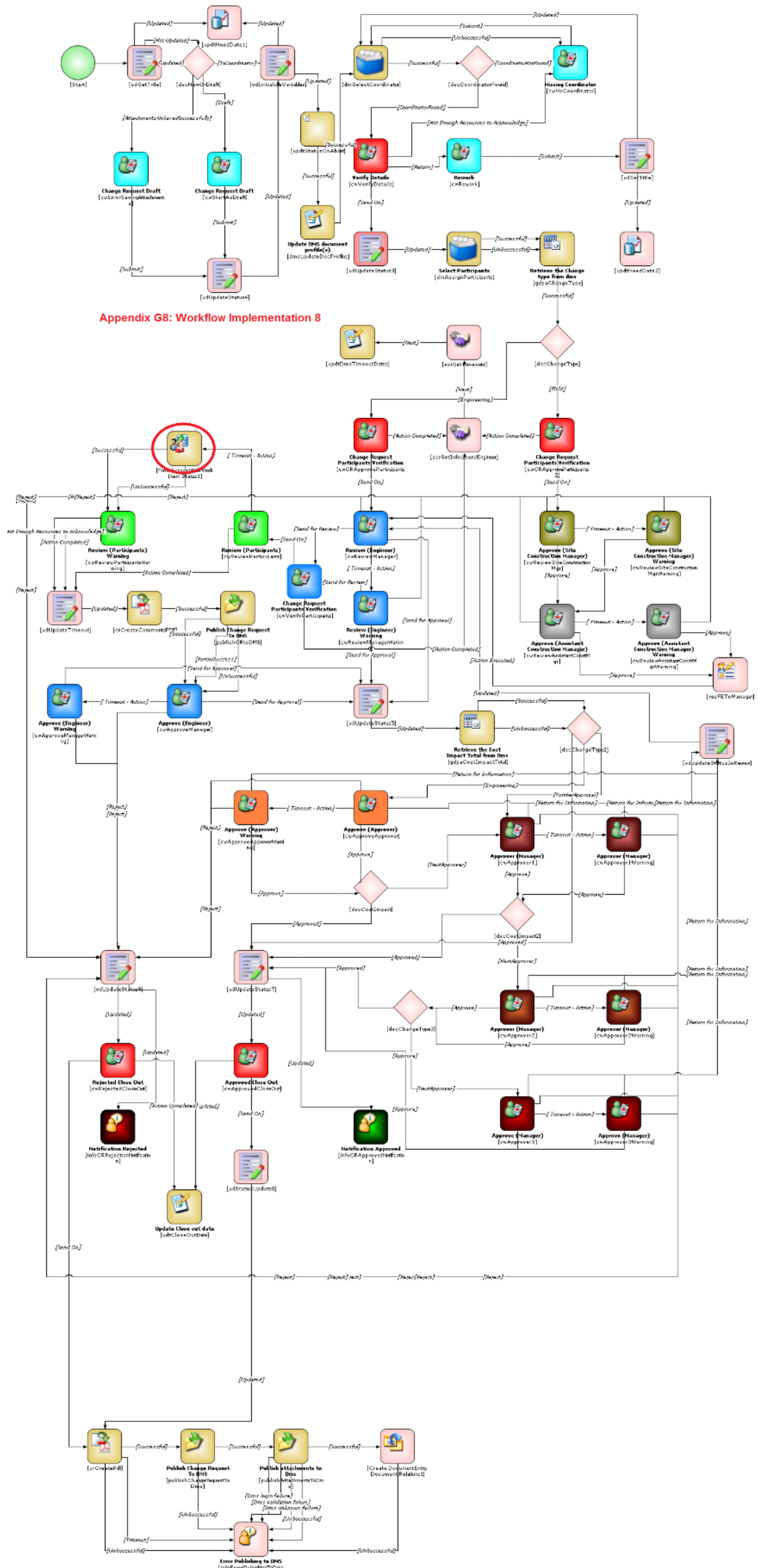


Figure F-8: Workflow Implementation 8

Appendix G: Two Versions of the Benchmark Workflow for Comparison

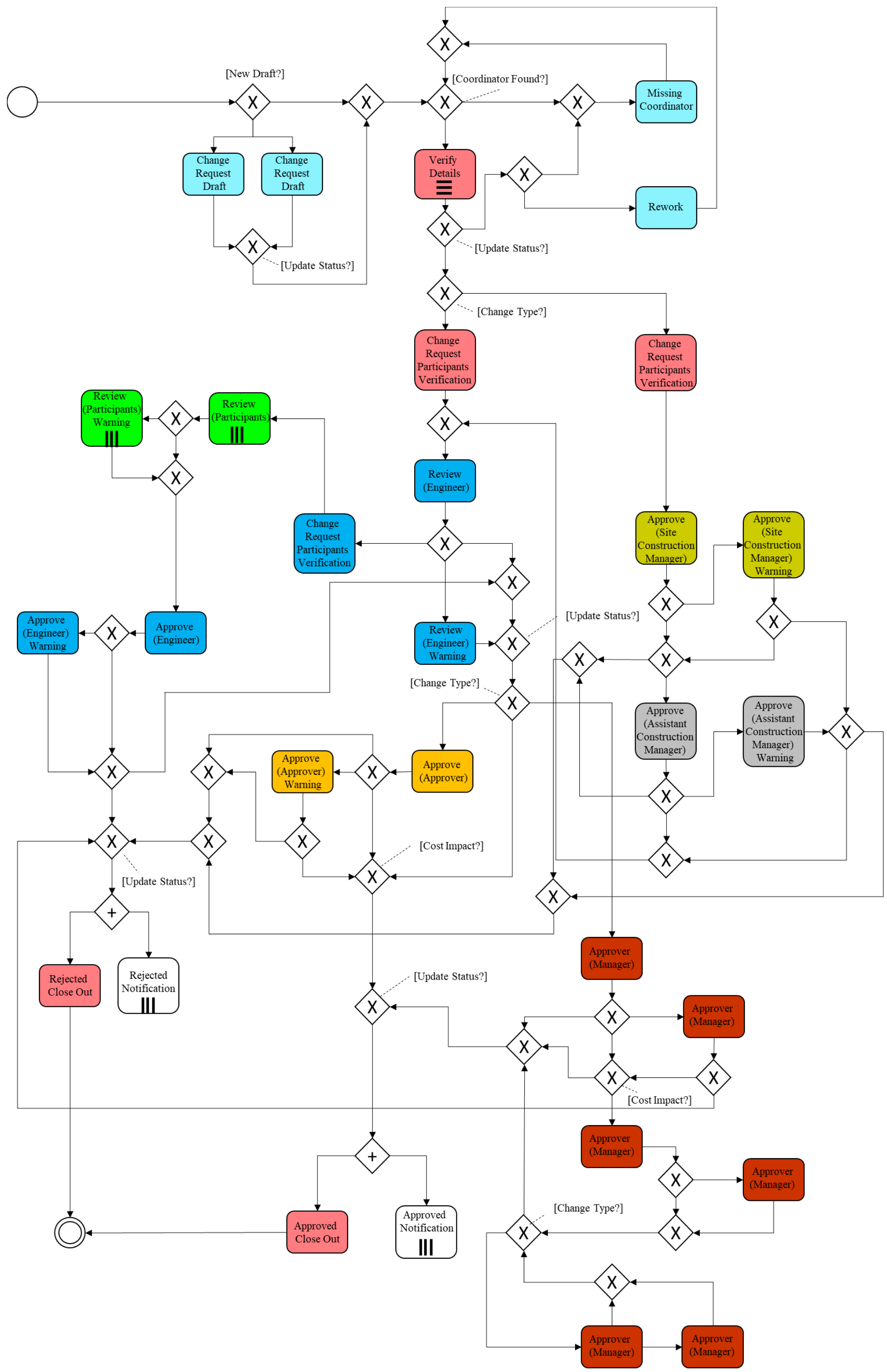


Figure G-1: Benchmark Workflow Revision 1 (BPMN)

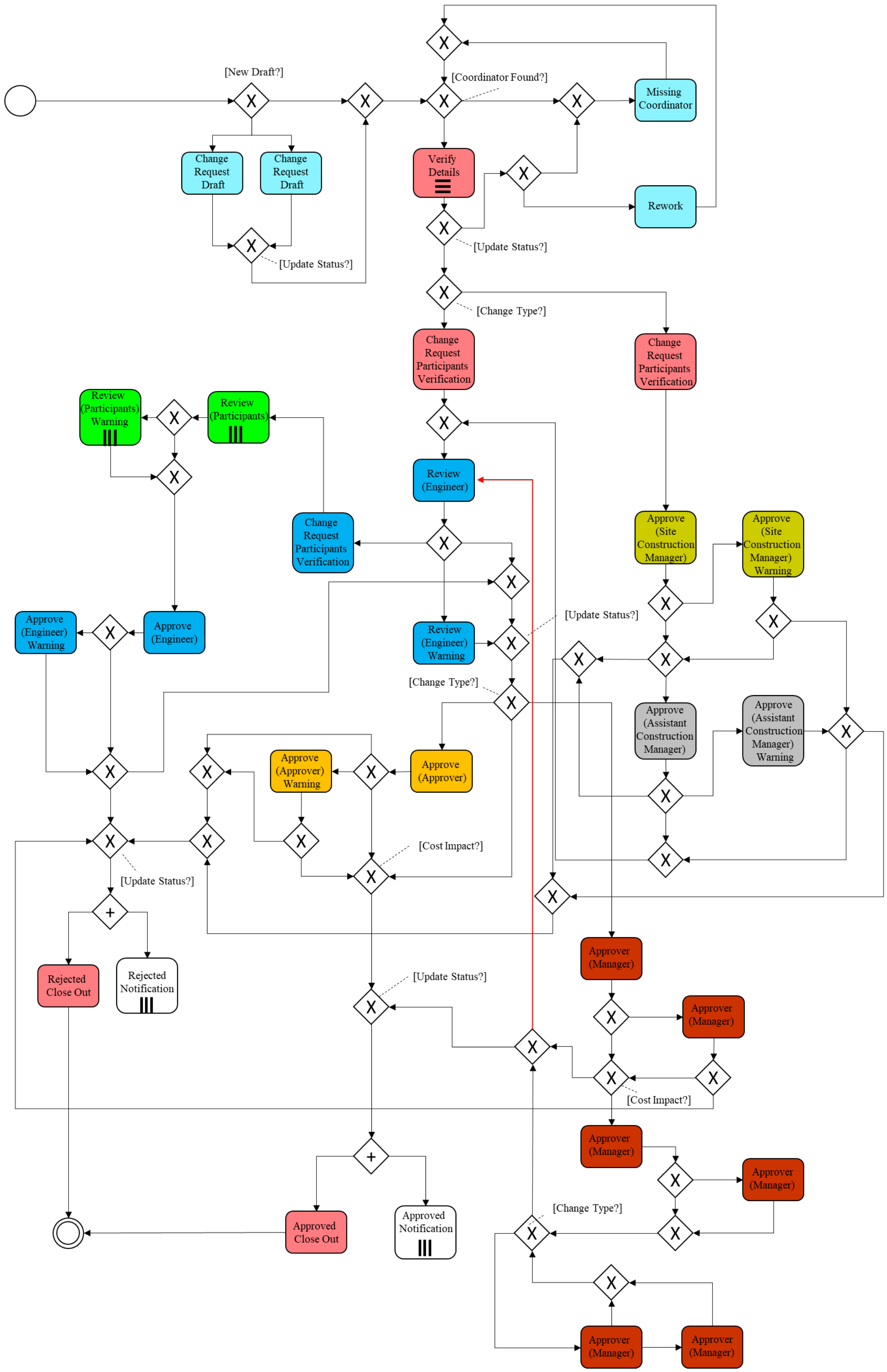


Figure G-2: Benchmark Workflow Revision 2 (BPMN)

Glossary

1. Accuracy: the number of correct (true) estimates over the number of total estimates
2. Adequate accuracy: application-specific accuracy that is driven by the cost of false positives and negatives that must be outweighed by the savings from automation to make the approach useful in practice
3. Artificial Intelligence (AI): the field of computer systems able to perform tasks that normally require human intelligence
4. Audit: an evaluation or examination by a person or group of people
5. Automation: the technology by which a process or procedure is performed with minimal human assistance
6. Automated compliance checking (ACC): the field of checking a text or semantic compliance of a model or document against geometric rules or regulations
7. Benchmark workflow: a workflow that is a standard or a target that a company attempts to pursue when implementing a process
8. Benchmarking and Metrics (BM&M) program: a data collection program that accumulates project data on the level of implementation to measure and assess capital project performance and find the best practices (Shan et al., 2011)
9. Best Practice: the description of the best way of working based on the situation in hand; or a process or method that, when executed effectively, leads to enhanced project performance (Construction Industry Institute)
10. Business Process Model and Notation (BPMN): a graphical representation for specifying business processes
11. Capital project: a long-term, capital intensive investment project with a purpose to build upon, add to, or improve a capital asset
12. Change Request (CR): formal proposal for an alteration to a process, product, or its components
13. Compliance: successfully fulfilling the guidelines (typically compulsory guidelines)
14. Compliance audit: an evaluation or examination of the policies and/or procedures of a company to check if they adhere to internal or regulatory guidelines
15. Compulsory guideline: a guideline which involves laws, policies, and codes which have the force of law
16. Computer vision: a field that deals with computers to gain high-level understanding from digital images or videos

17. Conformance: successfully fulfilling the guidelines (both compulsory and non-compulsory guidelines)
18. Conformance checking: a subfield (or function) of process mining which compares and evaluates a discovered workflow or event log against a benchmark workflow
19. Conformance component: a component of the practice conformance model (i.e., documents, workflows, people, and actions)
20. Conformance element: a tangible and measurable element of the practice conformance model (e.g., "process flowchart," "schedule/timeline")
21. Conformance map: a visual representation of results from the conformance measurement process which highlights conformances and/or non-conformances
22. Construction audit: an evaluation or examination of the costs incurred and actions taken for a specific construction project
23. Control logic: a key part (of a software program) that controls the operations of the program
24. Data mining: a field of examining large databases in order to generate new information
25. Discovered workflow: a visual workflow that is derived from the entire or a part of an event log by process discovery function
26. Document conformance: Conformance of a document related conformance element to practice guidelines
27. Document Management System (DMS): an electronic system to store, track, and manage documents for organization and accessibility
28. Effort to understand: the perceived effort that is required to understand a conformance element
29. Event log: a sequentially recorded collection of activities with corresponding case ID and timestamps, etc.
30. Event log preprocessing: a module to enhance the quality of a raw event log data in a format that can be understood by humans and machines
31. Financial audit: an evaluation and examination to increase the credibility of financial statements and reduce risks for stakeholders
32. First-party audit: an internal audit conducted by auditors who are employed by the company being audited but who have no interest in the audit results of the area being audited
33. Flowchart: a visual diagram designed to outline the flow typically of a process
34. Footprint-based method: one of the methods of the conformance checking field which compares two sets of footprints (direct relationships between two activities) derived from an event log or discovered workflow and a benchmark workflow

35. Framework: an entity between a "model" and a "method." A framework is, or contains, a (not completely detailed) structure or system for the realization of a defined result/goal. Many frameworks comprise one or more models, based on the modelling techniques mentioned above and often based on (best) practices. Compared with methods, frameworks give the users much more freedom regarding the (partial or entire) use of the framework
36. Frequency of use: the perceived frequency that a conformance element would appear in practices or guidelines
37. Full automation: automation (the term is used to compare with semi-automation)
38. Functional requirement: a description of the service that a framework or a toolset (software package) must offer
39. Image mining: a field examining large image databases in order to generate new information (e.g., machine vision, image processing, and image retrieval)
40. Impact: the degree to which an element may indicate conformance to a practice or a guideline
41. Information type: a category of elements from documents (e.g., practice conformance elements)
42. Keyword extraction: a module of extracting the most relevant words from texts
43. Lemmatization: a function of returning the base or dictionary form of a word
44. Level of importance: perceived importance of a conformance element
45. Matching: identical form (semantically or phonetically)
46. Machine learning: a subfield of the AI field which an application provides machines the ability to automatically learn and improve from experience without being explicitly programmed
47. Manual framework: a framework that does not involve any automation (e.g., Manual (WM))
48. Method: a systematic approach to achieve a specific result or goal, and offers a description in a cohesive and (scientific) consistent way of the approach that leads to the desired result/ goal
49. Model: the presentation in schematic form, often in a simplified way, of an existing or future state or situation
50. Module: each of a set of standardized parts that can be used to construct a framework
51. Natural Language Processing (NLP): a subfield of the AI field that analyzes, understands, and generates the languages that humans use naturally in order to interface with machines in both written and spoken contexts
52. Non-compulsory guideline: a guideline that involves best practices, benchmark processes, and/or standard procedures which do not have the force of law

53. Operational audit: an evaluation and examination of goals, planning processes, procedures, and results of the operations of a business
54. Practice: the description of the way in which professionals work within their profession, in order to carry out a specific task
55. Practice conformance: conformance to a practice or a guideline
56. Practice conformance model: a representation that describes practice conformance with conformance components, measurement methods, and conformance elements
57. Practice guideline: a manual indicating how to perform (compulsory/non-compulsory guideline)
58. Practice guideline document: an electronic- or paper-form document of practice guideline which usually is in text and/or image form
59. Process: a series of actions or steps taken in order to achieve a particular end
60. Process mining: a field to discover, monitor, control, and improve a process by extracting knowledge from an event log
61. RASCI (Responsible, Accountable, Support, Consulted, Informed): a matrix of decision-making authorities for all the activities undertaken in an organization
62. Related: associated with
63. Second party-audit: an external audit performed by a customer or contracted company on behalf of a customer
64. Semi-automation: a process or procedure that is performed by the combined activities of human and machine with both human and machine steps
65. Semi-automated framework: a framework that includes semi-automation
66. Significant difference: a statistical difference detected by hypothesis testing where the null hypothesis is rejected (e.g., 95% confidence level)
67. Software package: a set of software that fulfills a specific function
68. Stemming: a function of retaining the base meanings of a word but remove the last few characters
69. Stop word: a word which is generally the most common word in a language that is filtered out before processing the language data (e.g., preposition, conjunction, determiners)
70. Text extraction: a module of extracting text information from documents so that they are machine-readable
71. Text matching: a module of matching texts from documents
72. Text mining: a field of examining large text databases in order to generate new information

73. TF-IDF (Term Frequency-Inverse Document Frequency): numerical statistics used for extracting keywords by ignoring unrelated words
74. Third-party audit: an audit performed by an audit organization independent of the customer-company relationship and is free of conflict of interest.
75. Token-based method: one of the methods of the conformance checking field which replays a token (representation dot) that are produced, consumed, missing or remaining on a benchmark workflow
76. Tokenization: a process of breaking a stream of text up into words, phrases, symbols, or other meaningful elements called tokens in the text mining field
77. Toolset: a set of software tools
78. Validation: a process of checking whether a specification captures the needs of a customer
79. Verification: a process of checking whether a software package meets the specification
80. Whitelist violation: a tolerated violation that is considered as conformance (Celonis)
81. Workflow: a sequential flow of tasks to achieve some results, typically visually represented
82. Workflow conformance: Conformance of a workflow related conformance element to practice guidelines (e.g., benchmark workflow)
83. Workflow Management System (WMS): an electronic system that automatically routes the data to the predetermined task